Vânia Maria Gomes de Almeida

NEW METHODS FOR HEMODYNAMIC EVALUATION: A multi-parametric approach

Dissertation submitted for the degree of Doctor in Physics
(Technological Physics)

September 2013

· U C ·

UNIVERSIDADE DE COIMBRA

FACULTY OF SCIENCES AND TECHNOLOGY
DEPARTMENT OF PHYSICS

NEW METHODS FOR HEMODYNAMIC EVALUATION:

A multi-parametric approach

Scientific Supervisor:
Doutor João Manuel Rendeiro Cardoso

Co-Supervisor:
Professor Doutor Carlos Manuel Bolota Alexandre Correia

Thesis submitted in accordance with the requirements of the
University of Coimbra for the degree of Doctor in Physics

(Technological Physics)

VÂNIA MARIA GOMES DE ALMEIDA

SEPTEMBER 2013

This work is funded by National funds through Fundação para a Ciência e a Tecnologia - FCT - through the scholarship SFRH/BD/61356/2009 and by FEDER, through the Programa Operacional Factores de Competitividade - COMPETE - in the scope of project PTDC/SAU-BEB/100650/2008 .

Aos meus pais

Agradecimentos

Gostava de agradecer, a todos aqueles que ao longo dos últimos anos tornaram este trabalho possível. Ao meu orientador, Doutor João Cardoso, pela presença constante e disponibilidade, que foram essenciais no decorrer deste trabalho. Ao Prof. Doutor Carlos Correia por ter desde cedo acreditado em mim, e ter sido fonte de inspiração ao longo do meu percurso académico. Um agradecimento especial ao Prof. Doutor Requicha Ferreira, que apesar de não ser oficialmente orientador, teve um papel fundamental durante todo este percurso.

A nível institucional gostaria de agradecer todo o apoio financeiro que recebi por parte da Fundação para a Ciência e Tecnologia (FCT) e do Centro de Instrumentação da Universidade de Coimbra. Agradeço igualmente o apoio do Centro Hospitalar e Universitário da Universidade de Coimbra (CHUC), onde foram feitos os testes clínicos, principalmente o apoio prestado pelo Doutor Mariano Pegô e pelo Departamento de Cardiologia deste hospital.

A todos os colegas de laboratório, pelo companheirismo demonstrado ao longo dos últimos anos, em especial à Catarina Pereira e à Tânia Pereira, que comigo desenvolveram trabalho no campo da hemodinâmica. Ao João André Vieira, ao João Borba e ao Pedro Santos devo um agradecimento especial, já que contribuíram activamente para este trabalho, tendo sido um prazer ser supervisora dos vossos projectos de mestrado. Aos que nos últimos tempos têm sido óptima companhia, Mariana Sequeira e Pedro Vaz. Aos ex-colegas de laboratório que por aqui passaram, e já voaram para novos rumos, mas que também me acompanharam nesta caminhada, Ricardo Oliveira e Sónia Semedo, e em especial à minha amiga Edite Figueiras, para quem não tenho palavras que cheguem para agradecer a ajuda sempre presente, mesmo nos últimos tempos longinquamente desde a Finlândia, muito obrigado por tudo.

Finalmente, o agradecimento mais importante, aos meus pais, a quem devo tudo isto, já que me proporcionaram as melhores condições que eu alguma vez poderia ter. Na realidade podiam ser os melhores pais do mundo fazendo muito menos. Ao meu irmão que tem estado sempre por perto. Para além deles não posso deixar de esquecer aquele que a par dos meus avós, os que já cá não estão e a minha avó Emília, sempre olhou por mim, o meu tio Norberto Gomes da Costa, que desde cedo foi uma fonte de inspiração, fazendo-me acreditar que a adversidade não existe. Luís Costa e Rui Costa o agradecimento ao pai também se estende a vocês.

Por fim, um agradecimento muito especial ao Filipe Ferreira pelo carinho, pela partilha das dificuldades, pela preocupação, mas acima de tudo por todos os momentos bons, que sempre foram iluminando este caminho com muita felicidade.

Abstract

Cardiovascular diseases (CVD) are the leading cause of death worldwide and arterial stiffness, in particular, is considered a major determinant of risk. This thesis meets its motivation in the foreseeable impact that an accurate, non-invasive and easy-to-use instrument for hemodynamic condition assessment could impart to the diagnosis and follow-up of CVD. Technically, the approach to this question is based upon on the belief that the Arterial Pressure Waveform (APW) conveys invaluable information about the arterial stiffness which can be accurately evaluated by means of moderate-cost instrumentation. The lack of efficient APW analysis approaches, as well as the growing interest in clinical decision support systems, brought extra motivation to this work. The main goal along this work was the development of an electronic platform dedicated to arterial stiffness assessment. Its full implementation included four main steps: (1) a multi-modular platform designed to incorporate several clinical data modules, namely the APW module designed to accurately retrieve pulse morphology; (2) an algorithmic tool capable of extracting relevant information from the APW signal; (3) sets of validation tests used in the assessment of probe repeatability and reproducibility and; (4) machine learning algorithms capable of dealing with vectorized features. The first trials of the APW module consisted in the assessment of its piezoelectric front end performance (interface between the patient and the probe itself) in a special purpose test bench, where a root mean square error (RMSE) of $1.8 \pm 0.22\%$ was observed using artificial cardiac-like pressure waves. The probe was incorporated in a collar and tested using a specially developed database built from 217 subjects, organized in groups, with different characteristics. It has been demonstrated that the identification of different time and amplitude parameterizations for each group is possible, as well the identification of physiological changes in a group of subjects, after a ”stent” implementation. The validation tests showed reproducible and repeatable results whenever data were obtained by well-trained operators. The use of data mining techniques demonstrated reliable results, in particular, classification algorithms (*e.g.*, the Random Forest algorithm achieved an accuracy of 96.95%) and clustering methodologies, for which a good correlation between the high risk labels (retrieved from the implemented approach) and the Augmentation Index (AIx) was observed. Globally, results give clear positive indications concerning the use of the developed methodology (prototype and associated algorithmic tools) as an alternative to costly techniques, such as ultrasound or applanation tonometry, in clinical and in research environments.

Keywords: *biomedical instrumentation, multichannel data acquisition system, learning algorithms, cardiovascular diseases, arterial stiffness*

Resumo

As doenças cardiovasculares (DCV) são a principal causa de morte mundial, e a rigidez arterial, em particular, é considerada um importante factor de risco. Esta tese encontra a sua motivação no impacto previsível que um instrumento preciso, não invasivo e fácil de usar para avaliação da condicão hemodinâmica poderá dar ao diagnóstico e acompanhamento das DCV. Tecnicamente, a abordagem a esta questão baseia-se na crença de que a onda de pressão arterial (OPA) transmite informações valiosas sobre a rigidez arterial, e que esta pode ser estudada com precisão por instrumentação de custo moderado. A inexistência de abordagens de análise da OPA eficientes, bem como o crescente interesse em sistemas de apoio à decisão clínica, deram motivação extra a este trabalho. O principal objectivo deste trabalho foi o desenvolvimento de uma plataforma electrónica dedicada à avaliação da rigidez arterial. A sua implementação incluiu quatro etapas fundamentais: (1) uma plataforma multi-modular concebida para incorporar vários módulos clínicos, nomeadamente o módulo OPA projectado para reprodução da morfologia de pulso; (2) uma ferramenta algorítmica capaz de identificar informação importante na OPA; (3) testes de validação usados na avaliação da repetibilidade e reprodutibilidade da sonda e; (4) algoritmos de aprendizagem implementados para lidar com características vectorizadas. Os primeiros ensaios do módulo OPA consistiram na avaliação do desempenho da sua sonda piezoeléctrica (interface entre o paciente e o próprio módulo) em uma bancada de teste concebida para este propósito, onde um erro quadrático médio (RMSE) de $1.8 \pm 0.22\%$ foi observado, usando ondas de pressão sintetizadas. A sonda foi incorporada em um colar, e testada usando um banco de dados especialmente desenvolvido, construído a partir de 217 sujeitos, organizados em grupos, com diferentes características. Foi demonstrado que a identificação de diferentes parametrizações temporais e de amplitude para cada um dos grupos é possível, bem como a identificação de alterações fisiológicas em um grupo de sujeitos, após a implantação de um ”stent”. Os testes de validação mostraram resultados reprodutíveis e repetíveis, quando os dados foram obtidos por operadores bem treinados. O uso de técnicas de mineração de dados demonstrou resultados confiáveis, em particular, para os algoritmos de classificação (*p.ex.*, o algoritmo *Random Forest* alcançou uma precisão de 96.95%) e metodologias de agrupamento, para as quais foi obtida uma boa correlação entre as indicações de alto risco (obtidas usando a abordagem desenvolvida) e o Índice de Aumentação (AIx). Globalmente, os resultados dão bons indicadores, sobre o uso desta metodologia (protótipo e ferramentas algorítmicas associadas), em alternativa a técnicas dispendiosas, tais como técnicas de ultra-sons ou a tonometria de aplanação, em ambiente clínico e de investigação.

Keywords: *instrumentação biomédica, sistema de aquisição de dados multicanal, algoritmos de aprendizagem, doenças cardiovasculares, rigidez arterial*

List of Acronyms

AAMI	Association for the Advancement of Medical Instrumentation
ABPM	Ambulatory Blood Pressure Monitoring
ADM	Active Differentiator Mode
AIx	Augmentation Index
ANOVA	Analysis of Variance
ANN	Artificial Neural Network
APW	Arterial Pressure Waveform
ASSIGN	ASsessing cardiovascular risk using Scottish Intercollegiate Guidelines Network
AUC	Area Under Curve
BLR	Baseline Restorer
BP	Blood Pressure
BPT	Blood Pressure Treatment
CHD	Coronary Heart Disease
CAD	Coronary Artery Disease
CCII	Second Generation Current Conveyor
CMA	Charge Mode Amplifier
CMRR	Common-Mode Rejection Ratio
CRP	C-Reactive Protein
CV	Coefficient of Variation
CVD	Cardiovascular Diseases
DAQ	Data Acquisition
DAQP1	First Prototype Version
DAQP2	Second Prototype Version
DAQFP	Final Prototype
DBP	Diastolic Blood Pressure
DSP	Digital Signal Processor

DN	Dicrotic Notch
DP	Dicrotic Peak
DW	Dicrotic Wave
EM	Expected Maximization
ECG	Electrocardiogram
ESRD	End Stage Renal Disease
FH	Family History
FN	False Negative
FP	False Positive
FWHM	Full Width at Half Maximum
FRS	Framingham Risk Score
HDL	High-Density Lipoprotein
HEE	Human Expert Engineer
HR	Heart Rate
IMT	Intima-Media Thickness
LDL	Low-Density Lipoprotein
LpPLA2	Lipoprotein-associated Phospholipase
LVET	Left Ventricular Ejection Time
MLP	Multi-Layer Perceptron
NTC	Negative Temperature Coefficient
OTA	Operational Transconductance Amplifier
P*i*	Point of Inflection
PPG	Photoplethysmography
PROCAM	PROspective Cardiovascular Münster
PWV	Pulse Wave Velocity
RAAS	Renin-Angiotensin-Aldosterone System
RWTT	Reflected Wave Transit Time

PDF	Probability Density Function
PZ	Piezoelectric
RBF	Radial Basis Function
RMSE	Root Mean Square Error
RMSSD	Root Mean Square Successive Difference
ROC curve	Receiver Operating Characteristic curve
SBP	Systolic Blood Pressure
SD	Standard Deviation
SEVR	Sub-Endocardial Viability Ratio
SCORE	Systematic Coronary Risk Evaluation Project
SP	Systolic Peak
SWTT	Systolic Wave Transit Time
SMK	Smoking habits
SNR	Signal-to-Noise Ratio
SVM	Support Vector Machines
TGL	Triglycerides
TN	True Negative
TP	True Positive
UART	Universal Asynchronous Receiver/Transmitter
VMA	Voltage Mode Amplifier
WEKA	Waikato Environment for Knowledge Analysis
WRI	Wave Reflection Index
WHO	World Health Organization

Contents

Agradecimentos **v**

Abstract **vii**

Resumo **ix**

List of Acronyms **xi**

Contents **xv**

List of Figures **xxiii**

List of Tables **xxvii**

1 Introduction **1**

1.1 Motivation 2

1.2 Goals 3

1.2.1 Multichannel Biomedical Platform 3

1.2.2 Data Collection and Validation Tests 4

1.2.3 Hemodynamic Features 4

1.2.4 Clinical Assessment Support Systems 4

1.3 Structure 5

1.4 Scientific Contribution 6

2 Cardiovascular Diseases **9**

2.1 Motivation . . . 10

2.2 Risk Factors . . . 12

2.2.1 Non-Modifiable . . . 13

2.2.1.1 Ageing . . . 13

2.2.1.2 Genetics . . . 13

2.2.1.3 Socio-Economic Status . . . 13

2.2.2 Modifiable . . . 13

2.2.2.1 Obesity . . . 13

2.2.2.2 Blood Pressure . . . 14

2.2.2.3 Diabetes . . . 14

2.2.2.4 Smoking . . . 15

2.2.2.5 Biomarkers . . . 15

2.2.3 Prevalence in Portugal . . . 17

2.3 Diagnostic Tools . . . 18

2.3.1 Blood Pressure Monitoring . . . 18

2.3.2 Applanation Tonometry . . . 19

2.3.3 Photoplethysmography . . . 19

2.3.4 Optical Sensors . . . 19

2.3.5 Angiography . . . 19

2.3.6 Magnetic Resonance Imaging . . . 20

2.3.7 Ultrasounds . . . 20

2.4 Parameters used in Arterial Stiffness Assessment . . . 20

2.4.1 Pulse Wave Velocity . . . 21

2.4.2 Waveform Parameters . . . 21

2.4.3 Diameter Based Parameters . . . 22

2.5 Clinically Approved Devices . . . 24

2.5.1 Standardization of Measurements . . . 27

2.5.2 Accuracy Validation Studies . . . 27

2.6 Reference Values . . . 30

2.6.1 Blood Pressure . . . 30

2.6.2 Body Mass Index in Adults . . . 30
2.6.3 Dyslipidemia . . . 31
2.6.4 Augmentation Index . . . 31
2.6.5 Pulse Wave Velocity . . . 33
2.6.6 Waveform Analysis . . . 34
2.7 Risk Assessment . . . 34
2.7.1 Risk Scores . . . 35
2.7.2 Machine Learning Techniques . . . 35
2.7.2.1 Classification Techniques . . . 35
2.7.2.2 Clustering techniques . . . 36

3 Multichannel Biomedical Platform 37
3.1 Motivation . . . 38
3.2 System Architecture . . . 39
3.2.1 Power Supply Module . . . 39
3.2.1.1 Autonomy . . . 40
3.2.2 Bus . . . 41
3.2.3 Individual Models . . . 41
3.2.4 Synchronization . . . 43
3.2.5 Database . . . 44
3.3 Arterial Pressure Waveform Module . . . 45
3.3.1 Hardware Architecture . . . 46
3.3.1.1 Piezoelectric Sensors . . . 47
3.3.1.2 Amplifying Stage . . . 49
3.3.1.3 Peak Detector and Timer . . . 50
3.3.1.4 Integrator with Reset . . . 51
3.3.1.5 Baseline Restorer . . . 52
3.3.1.6 Breathing Circuit . . . 54
3.3.1.7 Microcontroller Circuit . . . 55
3.3.2 Firmware Development . . . 56
3.3.2.1 Pulse-by-pulse Control . . . 57

3.3.2.2 Simulink Implementation 59
3.3.2.3 *dsPIC* Programming 60
3.3.3 Probe Mechanical Design 61
3.4 Performance Tests 62
3.4.1 Spectral Bandwidth 62
3.4.2 Analogue Circuit Transfer Function 62
3.4.3 Input/Output Error 65
3.4.4 Pulse Delay 67
3.4.5 Optimal Mechanical Design 68
3.5 Test Bench Characterization 69
3.6 Conclusion 71
3.6.1 Manufacturing Costs 72

4 Data Collection and Processing **73**
4.1 Recruitment of Participants 74
4.1.1 First Trials (GroupF) 75
4.1.2 Repeatability Cohort (GroupR) 75
4.1.2.1 Trial 1 75
4.1.2.2 Trial 2 76
4.1.3 Control Cohort (GroupC) 77
4.1.4 Hypertensive Cohort (GroupH) 77
4.1.5 Angioplasty Cohort (GroupA) 77
4.1.6 Data Prediction Cohort (GroupD) 78
4.2 Patient Characteristics 78
4.3 Risk Factor Characteristics 80
4.4 Arterial Pressure Waveform Analysis 81
4.4.1 Pulse Waveform Modelling 82
4.4.2 Baseline Flutuations 83
4.4.3 Criteria for Pulse Flagging 83
4.4.4 Pulse Segmentation and Normalization 84
4.4.5 Pulse Contour Algorithm Implementation 85
4.4.5.1 Performance Evaluation 87
4.4.6 Pulse Waveform Features 89

5 Clinical Assessment Support Systems **93**

5.1 Motivation . . . 94

5.2 Feature Subset Selection . . . 95

5.3 Classification . . . 95

5.3.1 Single Classifiers . . . 95

5.3.1.1 J48 . . . 95

5.3.1.2 Random Forest . . . 96

5.3.1.3 RIPPER . . . 96

5.3.1.4 BayesNet . . . 96

5.3.1.5 Artificial Neural Networks . . . 97

5.3.2 Classifier Ensemble - Bagging and Voting . . . 97

5.4 Clustering . . . 98

5.4.1 *k*-Means . . . 98

5.4.2 Expectation-Maximization . . . 99

5.5 Risk Scores - Reference Method . . . 99

5.6 Methods . . . 100

5.6.1 Training . . . 100

5.6.2 Classifiers Performance . . . 101

5.6.2.1 Learning Curves . . . 102

5.6.2.2 Artificial Neural Network Performance . . . 103

5.6.3 Predictive Models . . . 103

5.6.4 Clustering . . . 103

5.6.4.1 Number of Clusters . . . 104

5.6.5 Risk Evaluation . . . 104

5.7 Results . . . 105

5.7.1 Ranked Feature List . . . 105

5.7.2 Classifiers Configuration . . . 105

5.7.2.1 Learning Curves . . . 105

5.7.2.2 Artificial Neural Networks . . . 107

5.7.3 Classifiers Selection . . . 110

5.7.4 Predictive Model Evaluation . . . 110

5.7.5 Comparison of Risk Scores Estimates . . . 112
5.7.6 PWA Parameters Evaluation . . . 116
5.7.7 Clustering Analysis . . . 117
5.7.7.1 GroupC . . . 118
5.7.7.2 GroupH . . . 121
5.7.7.3 Comparison of Groups . . . 123
5.8 Discussion . . . 124
5.8.1 Features Selection . . . 124
5.8.2 Classification Model . . . 125
5.8.3 Predictive Model . . . 125
5.8.4 Risk Scores . . . 126
5.8.5 Clustering Analysis . . . 127
5.9 Conclusion . . . 128

6 Validation Tests **129**

6.1 Motivation . . . 130
6.2 Methods . . . 130
6.2.1 Descriptive Statistics Analysis . . . 131
6.2.2 Normality Assessment . . . 131
6.2.3 Parametric and Non-Parametric Tests Application . . . 131
6.2.4 Exploration of Relationships Among Variables . . . 132
6.2.4.1 Pearson Product-Moment Correlation . . . 133
6.2.4.2 Spearman's Rank Order Correlation . . . 133
6.2.5 Comparison of Groups . . . 133
6.2.5.1 Independent Samples T-Test . . . 133
6.2.5.2 One-Way ANOVA . . . 134
6.2.5.3 Mann-Whitney U Test . . . 134
6.2.5.4 Kruskal-Wallis Test . . . 134
6.2.6 Intra and Inter-observer Variability . . . 134
6.3 Results . . . 135
6.3.1 Distribution of Values . . . 135

6.3.2 Intra-arterial *versus* Non-invasive Waveforms 137
6.3.3 Normality Assessment 139
6.3.4 Significance of Differences 139
6.3.4.1 Categorization by Month 139
6.3.4.2 Categorization by Operator 140
6.3.4.3 Categorization for the Left and Right Carotid Arteries 141
6.3.4.4 Categorization by Subject 142
6.3.5 Intra-observer Variability 146
6.3.6 Inter-observer Variability 147
6.3.7 Pulse-by-pulse Variability 148
6.4 Conclusion 149

7 Concluding Remarks **151**
7.1 General Conclusions 152
7.1.1 Future Research Lines 154

Appendices **157**
A Database Structure 158
B Simulink Configurations 161
C WEKA Toolkit 162
D Scores 165
D.1 Framingham Risk Score (FRS) 165
D.2 PROspective Cardiovascular Münster (PROCAM) 166
D.3 ASsessing cardiovascular risk using Scottish Intercollegiate Guidelines Network (ASSIGN) 167
D.4 Systematic Coronary Risk Evaluation Project (SCORE) 168
E Additional Statistical Data 169
E.1 Operators Significance 169
E.2 Sub-categorization by Subject 171
E.3 Trials Significance 176

References **179**

List of Figures

2.1 Structure of a medium-size elastic artery 11

3.1 Multichannel biomedical platform. 38

3.2 System architecture of the multichannel platform, where M1, M2, M2, Mn are the individual models, and PS is the power supply unit. 39

3.3 Battery power test module. 40

3.4 Discharge behaviour of the load acid battery with $V_{begin} = 12.5\ V$ and $V_{cut} = 11\ V$. . 41

3.5 Backplane bus configuration. 42

3.6 Detailed front panels of some of the modules: 1-PPG module, 2-ECG module, 3-PWV module, 4-NI 6210 module. 43

3.7 Cardiocheck main menu. 44

3.8 First and second prototype versions developed along this thesis. 45

3.9 The final version of the APW module prototype based on the *dsPIC*-based core, housed in a eurocard (100x160 mm^2) format. 47

3.10 Complete circuit with its main functional blocks enclosed in the shaded boxes. V_{Test} and V_{Out} are the voltages from which performance of the circuit is evaluated. V_d is a monitoring point for the circuits studied in 3.3.1.2. These voltages are outlined by circles. 48

3.11 PZ sensor characteristics: (a) *Thevénin* equivalent and (b) its frequency response, obtained by simulation. 49

3.12 The three main topologies for amplifying the sensor signal: (a) voltage or passive differentiation mode amplifier (VMA), (b) charge mode amplifier (CMA) and (c) active differentiator mode amplifier (ADM). 50

3.13 Amplitude and phase of the transferisk function of the circuits. Vertical axes are in decibels (dB) for amplitude and degrees for phase. 51

3.14 Equivalent electronic circuit for the *CCII*+. 53

3.15 BLR circuit (a) and its Orcad simulation for a trapezoidal signal (b), with (S ON, black line) and without (S OFF, grey line) restoration. 54

3.16 Breathing circuit implementation, where *Resp* is the respiratory signal. 55

3.17 Generic block diagram of the hardware connections in the APW DAQ board. The grey boxes indicate the main hardware signal blocks that communicate with *dsPIC*, and the blue blocks the hardware accessories required for the *dsPIC* operation. . . . 56

3.18 PZ waveform is represented in grey line and delayed PZ waveform in black line. The integrated waveform PZ int (integrated) represented in blue line was obtained from the delayed PZ waveform. The t_{min} represents the ideal delay value, t_u is the systolic upstroke time and t_d is the delay time used in the PZ waveform integration. 58

3.19 Simulink model used for C-code generation for *dsPIC* using Kerhuel blockset. . . . 59

3.20 Schematic of the stages of programming. (1) The model is generated in the Simulink platform and, then, is compiled in C code; (2) programming of the *dsPIC* using the MPLAB ICD2; (3) microcontroller real time data transmission through UART. . . 60

3.21 Mounting details of the probe elements shown in a transversal cut (a) and 3D exploded (b), views. The arrow in (a) indicates the externally applied forces and, (1) is the mushroom-shaped interface, (2) is the PZ disc sensor and (3) is the PCB frame. In (c) it is shown the upper view and probe dimensions, where (1) is the support PCB layer, (2) is the PZ disc and (3) the mushroom-shaped PVC (PolyVinyl Chloride) interface, and in (d) is shown a photo of the probe with collar. 61

3.22 Harmonic contents of the amplitude response of an APW segment, where is pointed the interest zone. Frequencies higher than 25 Hz are within in the noise plateau. . . 62

3.23 (a) Amplitude and (b) phase variation of the transfer functions of the ADM(4K), ADM(8K), VMA and CMA circuits. 63

3.24 Normalized input and output for a Type B cardiac pulse in the three circuit topologies, as given by the simulator (a) and measured in the circuit (b). Noise is only visible in the real data, plotted in (b). 66

3.25 In (a) the pulses acquired from both ADC channels are shown, where the delay between channels is clear and in (b) the pulse differentiation of both ADC channels is depicted. 68

3.26 RMSE for differently shaped sensor interfaces measured for a Type A waveform. In the first row, the three probe geometries are diagrammatically shown. The second row depicts the input and output acquired signals. The last row shows the RMSE distribution, demonstrating the superior consistency of the pointy probe. 69

3.27 Schematic representation of the test bench. The signal conditioning block encloses the circuit of Figure 3.10 and the device under test (DUT) is the probe presented in Figure 3.21. 70

3.28 Test bench characterization using a Gaussian shaped excitation. (a) Shows the mesh map of the probe signal along the tube, (b) depicts the excitation (black line) and probe signal at the identified tube position (48 *cm* from the ACT) (grey line). . . . 71

4.1 Database constitution by cohorts. 74

4.2 Risk factors distribution. 80

4.3 Distribution of subjects in study by age subsets. 80

4.4 Schematic representation of the pulse wave analysis process, for a Type A (a) and a Type C (b) waveform, where the APW features are identified. 81

4.5 Method adopted for baseline correction, where $y(n)$ and $y(n+1)$ are adjacent APW onset pulses. 83

4.6 Pulse wave analysis, (a) raw data; (b) baseline removal; (c) pulse segmentation and (d) the prominent points marked in a pulse during the segmentation process. 84

4.7 Flowchart representing the algorithm implementation. 86

4.8 Agreement between the SP and Pi obtained via PPIA and HEE, measured using regression plots (top) and Bland-Altman plots (bottom), (a) time information of SP and Pi, (b) relative amplitude of Pi. 89

5.1 Ranking of analysis features based on the average merit. 105

5.2 Classifier performance curves obtained by the variation of the number of subjects in the dataset. 106

5.3 Accuracy measured with increased subjects number. 107

5.4 Semi-log ROC curves for J48, Random Forest, JRIP amd BayesNet classifiers. . . 111

5.5 Boxplot of AIx distribution for classes A and B. The horizontal bold lines represent the medians, the box represent the interquartile ranges (50 % of the distribution) and the whiskers represent the range of values obtained for all subjects from GroupD in an overall set of 627 pulses. 112

5.6 Bland-Altman plots for the individual risk estimative by the FRS, PROCAM, ASSIGN and SCORE functions. 113

5.7 Correlation plots between transit time based parameters and all the considered risk scores (ASSIGN, PROCAM, SCORE and FSR). 117

5.8 Clusters performance obtained for EM (top) and *k*-means (bottom) algorithms using two clusters, Cluster 1=blue, Cluster 2=red. 118

5.9 RWTT and SWTT plot after *k*-means clustering for using three clusters. Blue=Cluster 1, red=Cluster 2 and Green=Cluster 3. 119

5.10 Scatter plots for: (a) RWTT and RWA, (b) LVET and SWTT, where the grey and black markers denote Cluster A and Cluster B, respectively. 121

6.1 Arrival time frequency distribution for SWTT, LVET and RWTT. 135

6.2 AIx values distribution for GroupC, GroupD and GroupH. 136

6.3 RMSSD successive differences measured for the SW, DW and RW: (a) time information (SWTT, LVET and RWTT); and (b) amplitude information (SWA, DWA, RWA). 137

6.4 Set of pulses detected invasively and non-invasively after and before the stent implementation. The Pi arrival points for each method are represented by red and blue circles, for pulses collected before and after angioplasty, respectively. 138

6.5 RWTT, SWTT and LVET comparison, before and after the angioplasty intervention. 138

6.6 Boxplot of the AIx categorized by subject and sub-categorized by month. 140

6.7 AIx values (right carotid artery) obtained for each subject and operator. 143

6.8 AIx values (left carotid artery) obtained for each subject and operator. 143

6.9 AIx values (right carotid artery) obtained for each subject and operator. 148

A.1 Cardiocheck *Acquisition* window: institution and patient settings. This window has the following fields: (A) institution select pop-up; (B) operator select pop-up and a field for notes; (C) place of acquisition select pop-up; (D) patient select pop-up and the corresponding information and notes; (E) proceed to signals and hardware window; (F) main menu button. 159

A.2 Cardiocheck *Search* menu: simulation of a patient search with "M" gender filtering condition. Fields in search window include: (A) entity selection; (B) filtering conditions; (C) search result list; (D) operations buttons; (E) main menu button. . . 159

C.1 The Weka Explorer: (a) choosing the *Explorer* interface and (b) visualization of the data (*Preprocess* panel). 162

C.2 Visualizing the result (a) of J48 classifier algorithm and, (b) its tree; (c) k-means clustering results, and (d) the cluster distributions. 163

C.3 *Visualize* menu. 164

D.1 FRS interface. 165

D.2 PROCAM interactive tool. 166

D.3 ASSIGN interactive tool. 167

D.4 SCORE interactive tool. 168

E.1 Bar plot for the SWTT measured at right carotid site. 171

E.2 Bar plot for the SWTT measured at left carotid site. 171

E.3 Bar plot for the RWTT measured at right carotid site. 172

E.4 Bar plot for the RWTT measured at left carotid site. 172

E.5 Bar plot for the LVET measured at right carotid site. 173

E.6 Bar plot for the LVET measured at left carotid site. 173

E.7 Bar plot for the DWA measured at right carotid site. 174

E.8 Bar plot for the DWA measured at left carotid site. 174

E.9 Bar plot for the RWA measured at right carotid site. 175

E.10 Bar plot for the RWA measured at left carotid site. 175

List of Tables

2.1 Overview of a set of studies on the association between wave reflections and clinical endpoints, sorted by date of publication, where n is the number of subjects. 23

2.2 Comparison of variability values obtained from different commercial devices. . . . 29

2.3 Definition and classification of hypertension. 30

2.4 AIx reference values for different populations 32

2.5 AIx reference values for healthy white European subjects, measured using SphygmoCor. 32

2.6 PWV reference values obtained using SphygmoCor system 33

2.7 Timing reference values for the RWTT, SWTT and LVET parameters. 34

3.1 Modules used in the multichannel platform. 42

3.2 Main characteristics of DAQ prototypes. 46

3.3 RMSE measurements for each of the three amplifying circuits, using cardiac - like synthesized pulses to reproduce the different classes described in the literature. Data are taken from simulator (PSpice) and measured in the circuit (real data). 66

3.4 Statistics of measurements depicted in Figure 3.26. 69

3.5 Table of costs. 72

4.1 The group variables. 75

4.2 GroupR characteristics for each one of the trials. 76

4.3 Popuplation characteristics for each analysed groups. 79

4.4 Parameters used in Equation 4.1, where $k = 1$ corresponds to the systole component, $k = 2$ corresponds to the reflection wave and $k = 3$ corresponds to the dicrotic wave, and w was used to sweep the range of interesting Pi situations. 82

4.5 APW features and abnormality criteria for pulse flagging. 84

4.6 Validation of algorithm performance (time) in comparison to HEE annotations. . . 88

4.7 Validation of algorithm performance (amplitude) in comparison to HEE annotations. 88

4.8 Ratios and indices used in the multi-parametric approach. 90

4.9 Pulse variance attributes. 91

5.1 Risk assessment tools (10 years term). 100

5.2 Confusion matrix of observed class *vs.* predicted class. 101

5.3 Parameters used in the assessment of the classifiers' performance. 102

5.4 MLP performance study for 1 hidden layer. 108

5.5 MLP performance study for 2 hidden layer. 109

5.6 RBF performance study, while manipulating clustering seed and the number of clusters. 109

5.7 Comparative classifier selection performance results. 110

5.8 Data obtained from class prediction analysis and the AIx values (for each subject). 111

5.9 Pearson's correlation matrix obtained between all variables and risk scores. 115

5.10 Average values for the 3-clusters features. 120

5.11 Cluster distributions for subjects in GroupH. 121

5.12 Spearman's correlation coefficients obtained for Cluster A (%). 122

5.13 Comparison of clusters from GroupH (Cluster A and Cluster B) and GroupC (Cluster 1, Cluster 2 and Cluster 3) using Kruskal-Wallis test, measured through χ^2 value. 123

6.1 Correlation coefficient ranges. 132

6.2 Summary of the techniques used in the exploration of the relationships among variables. 133

6.3 Descriptive data categorization by month. The *p-value* indicates the Mann-Whitney test significance value. 139

6.4 Descriptive data categorized by operator. The *p-value* indicates the Kruscal-Wallis test significance value. 141

6.5 Kruskal test significance for the left and right measurements. 141

6.6 Descriptive data categorization by subject for the left carotid artery. 144

6.7 Descriptive data categorization by subject for the right carotid artery. 145

6.8 Intra-observer mean differences $\pm$ 2 *SD* derived from the Bland-Altman plots ($n = 51$, obtained for the left carotid artery. 146

6.9 Intra-observer mean differences $\pm$ 2 *SD* derived from the Bland-Altman plots ($n = 51$, obtained for the right carotid artery. 146

6.10 Inter-observer mean differences $\pm$ 2 *SD* derived from the Bland-Altman plots ($n = 153$). 148

A.1 Cardiocheck main menu icons. 158
B.1 dsPIC MASTER 161
B.2 UART 2 Configuration 161
E.1 Descriptive data categorized by operator and sub-categorized by carotid site. . . . 170
E.2 Comparison of variability values obtained from commercial devices. 176

1

Introduction

1.1 Motivation

Over the last few years, the medical community has been discussing the exact relevance of vascular hemodynamics, with special focus on arterial stiffness [1–4]. It is well accepted by the scientific community that arterial stiffness is an early predictor of Cardiovascular Diseases (CVD) and, there are several studies that have stated its relation with diseases, such as: hypertension [5], coronary atherosclerosis [6], End Stage Renal Disease (ESRD) [7] and diabetes mellitus [8].

In what concerns the instrumental means available to researchers and clinicians, the literature shows that the real single issue behind the accurate assessment of arterial stiffness depends on the faithfully reproduction of the Arterial Pressure Waveform (APW), preferably using a non-invasive technique. Moreover, if this assessment could be performed with low cost devices, a significant impact on public health could be, herein, derived. Historically, the cuff sphygmomanometer, universally used by clinicians since the beginning of the 20th century, was the first device to quantify a part of the information contained in the APW, the Systolic Blood Pressure (SBP) and the Diastolic Blood Pressure (DBP), and conquered a (still) irreplaceable role in general clinical practice [9]. In the last years the major interest is with the central aortic pressure measurement [10], and the 24-h daily pressure monitoring [11]. On other hand, several physiologists and clinicians, started to pay increasing attention to other aspects of the information carried by the APW that can only be unfolded by its shape.

The history of arterial wave mechanic studies and applications is long. Being very early used by Chinese, Indian, Greek and Roman physicians in the diagnosis of several diseases [12]. In the last years, the separation of the incident from the reflected wave, and the analysis of parameters, such as, the Pulse Wave Velocity (PWV) [13] and the Augmentation Index (AIx) [14], assumed a enhanced relevance.

It is also accepted that the combination of several parameters (*e.g.*, biochemical and APW parameters) used in the clinical path of cardiovascular patients can provide a refined physiological status assessment. The recent emergence of computer-aided diagnosis technologies provide new insights for the development of innovative methods for professional health assistance in this area. Data mining techniques can be used in the recognition of relationships and patterns in huge amounts

of data, and so, being decisive in unveiling new biomedical and healthcare knowledge that can be used in new clinical decision support systems [15–17].

1.2 Goals

This project is of a strongly interdisciplinar nature incorporating two balanced components: engineering and medical sciences (as can be observed in the description, below). The hemodynamic analysis is supported by the combination of available clinical information, with the developed non-invasive instrumentation, which provides a detailed profile of the arterial pulse. To accomplish this, the key tool, is a multi-parameter approach based on data mining and machine learning algorithms capable to deal with a set of hemodynamic parameters.

1.2.1 Multichannel Biomedical Platform

This project finds its motivation in the foreseeable impact that an accurate, non-invasive and easy-to-use instrument for hemodynamic condition assessment could impart on the diagnosis and follow-up of the CVD. The APW conveys invaluable information about the condition of the cardiovascular system, and it can be accurately evaluated by means of moderate-cost instrumentation. Worldwide, the main specific devices, used in different experimental and clinical laboratories have been studied by other authors, such as the ones presented by Bruno Pannier [18]. The applanation tonometry [19] is a sophisticated and costly technique, that remains the preferred method for the APW analysis. Being this problem crucial, because these waves carry in their pattern information about all the mechanical factors that influence pressure and flow speed [20], it is important the development of precise and easy-to-use instrumentation.

A new system, capable of delivering this information was developed, being the piezoelectric (PZ) sensors studied and incorporated in a standalone module designed for APW monitoring. Two prototype versions were developed following these methodologies. However, the clinical trials evolved to a point, where it was considered the development of a multi-parametric integrated instrument able to include other technologies, in addition to this module, useful for the characterization of the patients.

1.2.2 Data Collection and Validation Tests

Generally, the validation of medical data acquisition systems requires the exploration of large databases. Therefore, the establishment of a database for this particular laboratory prototype is critical, using subjects not only with cardiac pathologies but, also, healthy subjects, covering a wide age range. A database for the APW module was established along this thesis by different group of subjects. Some of the subjects were recruited via health visitors in the Coimbra Hospital and University Centre (CHUC), Portugal, while others were obtained from the collaboration with volunteers at screening mobilizations. This database was essential for the assessment of the system repeatability (intra-operator variation) and reproducibility (inter-operator variation).

1.2.3 Hemodynamic Features

The selection of the most interesting hemodynamic indicators is an important task which has to take into account the relative relevance of each one. The AIx and the arterial wave reflections are classic themes of debate across the literature. However, this work also aims at providing the design of new variables that could be used in a multi-parametric approach to retrieve more information from the arterial blood pulse waveform, which have not yet been studied enough. So, the combination of carefully selected APW features should significantly improve the arterial blood pulse understanding, especially if compared with the traditional single parameter analysis, where the failure in the determination of one component, such as the inflection point, can jeopardize the whole assessment.

1.2.4 Clinical Assessment Support Systems

Most of the APW analysis frameworks result from the individually processing of each one of the hemodynamic parameters and not from the evaluating of inter-dependencies in the overall pulse morphology. One of the key contribution of this work is the application of machine learning algorithms to deal with the vectorized features, extracted from APW. This methodology provide a new insight into the analysis of a great number of parameters in such a way, that it is possible to emphasize the relationships on the hidden patterns, for the different datasets. Several approaches

have been used in many applications during the last few years, with interesting developments in medical fields. In this work were implemented clustering and classification algorithms that allowed the identification of different parameterizations for subjects of different groups.

1.3 Structure

This dissertation intends to provide a complete description of the developed multichannel electro-mechanical platform and its algorithmic tools, as well as the methodologies adopted to perform the system validation. The document can be systematized as follows:

Following this **Introduction** (Chapter 1), a short overview about CVD and the main risk factors associated to their development are presented in **Chapter 2**. Firstly, the major risk factors are shortly discussed, and then, a perspective analysis about the prevalence of CVD in Portugal is given. In this chapter are also discussed the common used methods in the cardiovascular pulse measurement. In addition, some of the most used CVD commercial devices are shortly discussed, as well as the most relevant parameters.

The **Chapter 3** describes the design and architecture of the three laboratory prototype developed versions. It is also presented the multi-modular system architecture, where the final APW prototype version was implemented. The remaining modules are shortly described.

Chapter 4 summarizes the main experimental measurements performed during pre-clinical tests. Experimental measurements were conducted in subjects during a mobilized screening week in the Department of Physics of the University of Coimbra, in collaboration with SCDSOS (Sudden Cardiac Death Screening of Risk Factors) study and, in pre-clinical trials at the Coimbra Hospital and University Centre (CHUC), among healthy, hypertensive and Coronary Artery Disease (CAD) subjects. The risk factors distribution is also briefly discussed. General information about signal processing techniques employed on the APW analysis are also presented.

The focus of the **Chapter 5** are the data mining techniques employed it. Data pre-processing, classification and clustering are the main detailed points. It also contains information about the suitable risk assessment tools (Framingham Risk Score (FRS), PROspective CArdiovascular

Münster (PROCAM), ASsessing cardiovascular risk using Scottish Intercollegiate Guidelines Network (ASSIGN) and Systematic COronary Risk Evaluation project (SCORE)).

Chapter 6 introduces some approaches concerning the evaluation of the inherent precision of the equipment, using general descriptive statistics, correlation techniques, and other relevant statistical tests that were conducted to find out statistical significant differences among groups. The performed tests included the characterization of the typical range of values for each group in the database, the comparison of invasive and of the non-invasive waveform measurements, the study of the intra-and inter-operator variances, as well as the analysis of the successive measurement differences.

The main conclusions and future potential improvements/developments of the proposed approach are discussed in **Chapter 7**.

1.4 Scientific Contribution

Throughout the work development some scientific publications have been published. These papers can be organized as: i) scientific international journals; ii) conference proceedings.

Papers in scientific international journals, with refereeing:

1. VG Almeida, M Pego, T Pereira, HC Pereira, C Correia JMR Cardoso, Cardiovascular risk analysis by means of pulse morphology and clustering methodologies, *submitted*.

2. VG Almeida, J Vieira, P Santos, T Pereira, HC Pereira, C Correia, M Pego, JMR Cardoso, Machine Learning Techniques for Arterial Pressure Waveform Analysis. Journal of Personalized Medicine, v. 3, n. 2, p. 82-101, 2013.

3. VG Almeida, HC Pereira, T Pereira, E Figueiras, E Borges, JMR Cardoso, C Correia, Piezoelectric probe for pressure waveform estimation in flexible tubes and its application to the cardiovascular system. Sensors and Actuators A: Physical, v. 169, n. 1, p. 217-226, 2011.

Papers in conference proceedings, with refereeing:

1. VG Almeida, J Borba, T Pereira, HC Pereira, J Cardoso, C Correia, Indices and repeatability tests of cardiovascular function performed on the arterial distension waveform - Case Study: Angiography intervention. In: INTERNATIONAL CONFERENCE ON BIO-INSPIRED SYSTEMS AND SIGNAL PROCESSING, 2013. Barcelona, Spain. Proceedings of the International Conference on Bio-inspired Systems and Signal Processing. SciTePress - Science and and Technology Publications, 2013, p. 138-144.

2. VG Almeida, J Borba, T Pereira, HC Pereira, J Cardoso, C Correia, Data Mining based methodologies for cardiac risk ptterns identification. In: INTERNATIONAL CONFERENCE ON BIOINFORMATICS MODELS, METHODS AND ALGORITHMS, 2013. Barcelona, Spain. Proceedings of the International Conference on Bioinformatics Models, Methods and Algorithms. SciTePress - Science and and Technology Publications, 2013, p. 127-133.

3. VG Almeida, LFR Ferreira, C Correia, Hemodynamic parameters assessment. In: 2012 IEEE 2ND PORTUGUESE MEETING IN BIOENGINEERING (ENBENG), 2012. Coimbra, Portugal. 2012 IEEE 2nd Portuguese Meeting in Bioengineering (ENBENG). IEEE, 2012, p. 1-6.

4. VG Almeida, J Borba, TM Pereira, HC Pereira, JMR Cardoso, C Correia, Validation of a waveform delineator device for cardiac studies: Repeatability and data mining analysis. In: 2012 IEEE 2ND PORTUGUESE MEETING IN BIOENGINEERING (ENBENG), 2012. Coimbra, Portugal. 2012 IEEE 2nd Portuguese Meeting in Bioengineering (ENBENG). IEEE, 2012, p. 1-4.

5. VG Almeida, T Pereira, E Borges, JMR Cardoso, C Correia, HC Pereira, 2011. A real time cardiac monitoring system - Arterial Pressure Waveform Capture and Analysis. In: INTERNATIONAL CONFERENCE ON PERVASIVE AND EMBEDDED COMPUTING AND COMMUNICATION SYSTEMS, 2011. Vilamoura, Algarve, Portugal. Proceedings of

the 1st International Conference on Pervasive and Embedded Computing and Communication Systems. SciTePress - Science and and Technology Publications, 2011, p. 83-90.

6. VG Almeida, TM Pereira, HC Pereira, JMR Cardoso, C Correia, C., A cardiac signal monitoring and processing system. In: 2011 1ST PORTUGUESE MEETING IN BIOENGINEERING - THE CHALLENGE OF THE XXI CENTURY (ENBENG), 2011. Lisbon, Portugal. 1st Portuguese Biomedical Engineering Meeting. IEEE, 2011, p. 1-4.

7. V Almeida, P Santos, E Figueiras, E Borges, T Pereira, J Cardoso, C Correia C., 2011. Hemodynamic features extraction from a new arterial pressure waveform probe. In: INTERNATIONAL CONFERENCE ON BIO-INSPIRED SYSTEMS AND SIGNAL PROCESSING, 2011. Rome, Italy. Proceedings of the International Conference on Bio-inspired Systems and Signal Processing. SciTePress - Science and and Technology Publications, 2011, p. 195-200.

8. V Almeida, T Pereira, E Borges, E Figueiras, J Cardoso, C Correia, HC Pereira, JL Malaquias and J B Simes, 2010. Synthesized cardiac waveform in the evaluation of augmentation index algorithms. In: 3RD INTERNATIONAL JOINT CONFERENCE ON BIOMEDICAL ENGINEERING SYSTEMS AND TECHNOLOGIES (BIOSIGNALS 2010), 2010. Valencia. Proceedings of Conference on Biomedical Engineering Systems and Technologies. 2010.

2

Cardiovascular Diseases

This chapter presents some of the more important physiological aspects of CVD. Along the sections, the most relevant risk factors are summarized and several diagnostic techniques are discussed. The more used commercial devices are presented and, the most relevant parameters used in research and clinical practise detailed. Nowadays, the concern to change the current healthcare paradigm, from reactive towards preventive, justifies the growing interest in non-invasive techniques and new tools for the management of the CVD risk assessment.

2.1 Motivation

According to the World Health Organization (WHO), CVD represent 31 % of all global deaths [21], and are the leading cause of death worldwide. As societies become more developed and the standard of living improves, CVD increase in prevalence due to the life expectancy growing. However, the improvements in living standards, are responsible by the decline of the mortality values, owing to increasing knowledge and treatment of risk factors. This means that, although CVD mortality rates are now decreasing in developed countries, the problem continues to escalate globally because rates in developing regions are increasing.

Arterial stiffness and pressure wave reflection have gained increased attention of the clinical community and, the prognostic value of these novel parameters and their association with cardiovascular disease has been the subject of many clinical studies [22, 23]. The arteries can be divided in four different groups: elastic arteries, medium muscular arteries, small arteries and arterioles. The elastic properties of the arteries change along the arterial tree, being the proximal arteries more elastic, and the distal arteries more stiff. Heterogeneity is caused by the molecular, cellular and histological structure of the arterial wall, which differs between the various parts of the arterial tree [19]. It is an important physiological and pathological factor that causes the pulse amplification along the arterial tree. This amplification occurs from central to distal arteries due to the occurrence of the wave reflections, that happens generally in peripheral arteries, since the reflection sites are closer to the peripheral arteries, than to the central arteries.

Aortic stiffness has an independent predictive value for fatal and non-fatal coronary events and fatal strokes in patients with various levels of cardiovascular risk [2, 24–26]. The aorta is the major vessel of interest, but measurements of the local carotid stiffness may also provide an important prognosis, since this artery is a frequent site of atheroma formation [2]. The carotid artery belongs to the group of elastic arteries, being the muscular media layer relatively thin and, therefore the thickening of the arterial wall is essentially caused by intimal thickening. The structure of a elastic artery wall is organized in three different layers, as presented in Figure 2.1. The innermost layer is the tunica interna (ou tunica intima), the middle layer is the tunica media and the outermost layer is the tunica externa (or tunica adventitia).

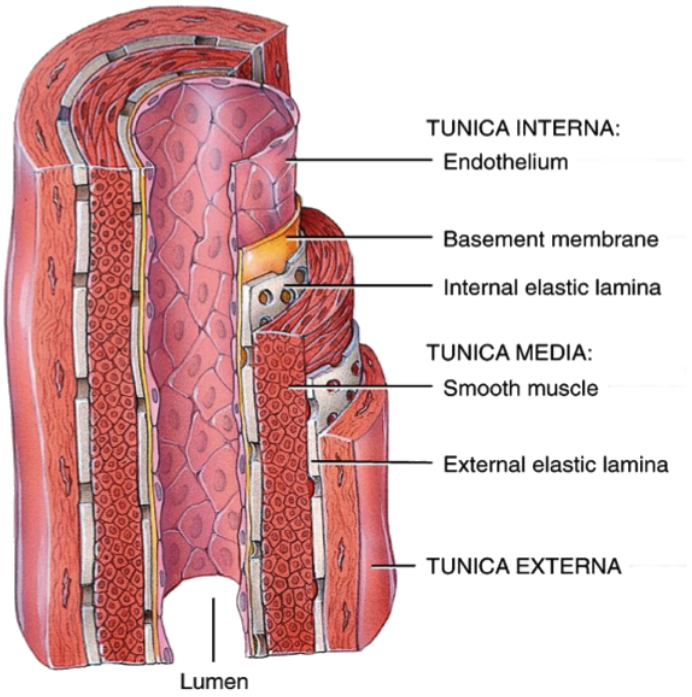

Figure 2.1: Structure of a medium-size elastic artery. Adapted from [27].

The increasing awareness of arterial stiffness [2, 28] started a new era of search for parameters capable of, directly or indirectly, quantifying its development. There are several advantages of using non-invasive methods, over invasive measurements. For instance, non-invasive methods can be used in follow-up trials of populations free from symptomatic CVD, such as children or young adults. Furthermore, these measurements might be a tool for risk stratification in addition to the established risk factors, for high risk populations, aiming the disease prevention. It is possible to assess regional and local arterial stiffness directly, and non-invasively, at various sites along the arterial tree. On the other hand, the systemic arterial stiffness can only be estimated from models of the circulation [2].

Increased arterial stiffness, determined invasively, has been shown to predict a higher risk of atherosclerosis, which it is involved in several pathological processes that can lead to heart attacks (coronary heart disease) and strokes (cerebrovascular disease) [4, 29]. Atherosclerosis is responsible by the narrowing of the arterial wall, and it is a serious disease that affects many people [30]. The main situations that effect the normal blood flow are: the encroaching of the lumen, which originates localized stenosis, the formation of thrombus that narrows the arterial lumen, the stiffening of arteries that leads to the wave reflections occurrence in upstream vessels, and the encroaching of the media that causes the destruction and atrophy of structural elements, which can lead to

aneurysms formation. The major sites associated to its development are the carotid, the cerebral, the epicardial coronary, the femoral and the iliac arteries, as well as the infrarenal part of the abdominal aorta. The major risk factors associated to its development are: age, male gender, hypertension, hypercholesterolemia, smoking habits (SMK) and adiposity [31] (see Section 2.2).

A heart attack can occur when the blood flow to the heart is cut off, due to a thrombus on a ruptured atherosclerotic plaque. This causes a shortage in the supply of oxygen and nutrients, that can damage the heart muscle. [21]. The pathology of ischaemic stroke includes, besides thrombus formation in atherosclerotic cerebral blood vessel, small vessels in the brain [32]. A possible bleeding caused by rupture of a blood vessel, caused by an aneurysm, can also be a cause. Additionally, strokes can also be caused by a blood clot that may travel through the blood vessels to the brain, as a consequence of irregular heartbeat [21].

There are evidence in studies of Dernellis and Panaretou [33] and of Liao *et al.* [34], together with earlier studies, that arterial stiffness and hypertension interacting with each other in a bidirectional manner. Liao *et al.* hypothesised that greater arterial stiffness in large and medium-sized arteries represents a cumulative adverse impact of conventional risk factors on the arterial wall, and that arterial stiffness, together with its adverse impact on other target organs, such as the kidneys, contributes to the development of hypertension. There are also evidence that aortic stiffness and arterial pulse wave reflections are key determinants of elevated central systolic pressure and that they are associated with adverse cardiovascular outcomes [35].

2.2 Risk Factors

Risk factors can be divided in two general types, of non-modifiable and modifiable natures. In the first group parameters are included parameters, such as hypertension, SMK and elevated cholesterol levels, whereas the non-modifiable group includes parameters, such as age, socio-economic status and heredity. Although, the non-modifiable factors cannot be modified, their presence helps to identify subjects at greatest risk, enabling vigorous treatment of the risk factors that can be modified.

2.2.1 Non-Modifiable

2.2.1.1 Ageing

Ageing is associated with changes in a number of structural and functional properties of large arteries, including diameter, wall thickness, wall stiffness, and endothelial function [36]. Progressive Intima-Media Thickness (IMT) caused by age has been reported [37]. Other studies showed that the vascular ageing is accelerated by the coexistence of several cardiovascular risk factors [38,39], such as hypertension or diabetes. Thus, it is expectable the reduction of the life expectancy for a young person, if preventive measures are not adopted.

2.2.1.2 Genetics

The importance of the familial prevalence of early-onset CVD is not yet completely understood in clinical practice [40]. However, there is evidence of a strong heritability relation of many cardiovascular risk factors in premature CAD [41]. Therefore, it is important that the familial prevalence of atherosclerosis disease, diabetes mellitus or hyperlipidaemia should be systematically investigated in the first-degree relatives with early-onset CAD in men (before 55 years old), and, in women (before 65 years old).

2.2.1.3 Socio-Economic Status

Familial history is also dependent of the surround environment. Multiple prospective studies have shown that people with low socio-economic statues, *i.e.* low educational level, low income, low-status job or living in a poor residential area, present an increase in the CVD mortality risk [42,43].

2.2.2 Modifiable

2.2.2.1 Obesity

Obesity has not been considered an independent risk factor for CAD development. However, fat, and in particular intra-abdominal visceral fat, is associated with the increase of the free fatty acids

secretion, hyperinsulinaemia, insulin resistance, hypertension and dyslipidaemia [44]. Abdominal obesity is also considered an independent risk factor for the development of atherosclerosis, namely in obese children [45]. The relationship between the obesity and aortic stiffness was already demonstrated [46], as well as the association of the abdominal fatty deposits with inflammatory markers, namely the C-Reactive Protein (CRP) [47].

2.2.2.2 Blood Pressure

High Blood Pressure (BP) is a risk factor for Coronary Heart Disease (CHD), heart failure, peripheral vascular disease, and renal failure in both men and women [48]. The amount of blood pumped by the heart, and the size and condition of the arteries are determinant in the BP [40]. However, many other factors can affect BP including the condition of kidneys and the levels of various hormones in the body. Some people experience essential hypertension, that is a type that has no identifiable cause, while others present secondary hypertension, that is caused by an external physiological condition, such as kidney problems or certain medicines.

There are evidences that BP values are associated with the increase of the relative risk of CVD [49, 50]. Recently, more emphasis has been assigned to the role of Pulse Pressure (PP) and isolated SBP in the identification of individuals at high risk. As a matter of fact, it is important to raise awareness regarding hypertension, the leading cause of mortality globally, and to renew efforts to reduce sodium intake and hypertension control [51].

2.2.2.3 Diabetes

Prevention of CVD in diabetic patients follows the same general principles as in healthy subjects, but with special concern with the glucose monitoring [40]. Several studies have demonstrated that diabetic patients have an increased risk, when compared with healthy participants [52, 53].

Diabetes are classified in accordance with the different clinical stages of hyperglycaemia. Four main categories have been identified in literature: type 1, type 2, other specific types, and gestational diabetes. Subjects with diabetes type 1 are characterized by a deficiency in insulin levels due to destructive lesions of pancreatic B-cells, which typically occurs for young subjects. Type 2 diabetes

are caused by the combination of the decrease in insulin secretion and sensitivity. Its prevalence comprises 90 % of adults with diabetes, and it is typically developed, after middle age, in obese and physically inactive patients. Gestational diabetes arises from glucose perturbations that occur during pregnancy, but it generally disappears after delivery. However, females with gestational diabetes have higher probability to develop diabetes over time [54].

2.2.2.4 Smoking

Cigarette smoking is the most important modifiable risk factor. It is responsible for 50 % of all avoidable deaths in smokers, half of these due to CVD [40]. Smoking is associated with increased risk of all types of CVD diseases. Although the rate of smoking is declining in Europe, it is still very common among individuals who have received little education; and widening education-related inequalities in smoking-cessation rates have been observed in many European countries in recent years [55]. Changing smoking behaviour is one of the most effective and cost effective things that can allow to improve CVD health, since several of the advantages are almost immediate, in spite of others take more time [51, 56].

2.2.2.5 Biomarkers

Established biomarkers associated to cardiovascular risk include dyslipidemia and glucose levels. However, the number of potential novel markers is expanding yearly, each time the possible candidates pass through the grading of clinical evidence. Novel biomarkers included: inflammatory (CRP, fibrinogen) and thrombotic markers (homocysteine), lipoprotein-associated phospholipase (LpPLA2) or the Renin-Angiotensin-Aldosterone System (RAAS). A short overview of these parameters is presented bellow:

- **Cholesterol**

  Several studies have established the crucial role of dyslipidaemia, namely the hypercholesterolaemia in the development of CVD [51, 57]. The evidence that the reduction of the plasma cholesterol contributes to decrease the CVD risk is strongly accepted. In blood plasma, the cholesterol is bound to various proteins (apoproteins) to form lipoproteins.

The most important types are the Low-Density Lipoproteins (LDLs) and the High-Density Lipoproteins (HDLs). Most of the cholesterol in blood plasma is normally carried in LDL and, over a wide range of cholesterol concentrations, it is strongly and positively associated with CVD risk [58]. Epidemiological studies, as well as trials with angiographic or clinical endpoints reinforce that the reduction of LDL cholesterol must be of prime concern to adopt during CVD prevention [40]. Every 1.0 $mmol/L$ reduction in LDL cholesterol could correspond to $20-25$ % of risk reduction in CVD mortality and non-fatal myocardial infarction [59]. Lower HDL concentrations are strongly associated with higher CVD risk levels. Several studies have demonstrated that the HDL cholesterol can contribute substantially to the risk prediction [40, 60]. There is also evidence that the combination of moderately elevated triglycerides (TGL) and low concentrations of HDL cholesterol is a very common situation in high-risk patients with type 2 diabetes, abdominal obesity, insulin resistance, or for those who are physically inactive [61].

- **Tryglicerides**

  The role of the plasma TGL, as a risk factor for CVD, for both men and women, in the general population, is being discussed for many years [62]. However, its effect is attenuated by the adjustment of other factors, namely the HDL cholesterol [40, 63].

- **High sensitive C-reactive protein**

  High-sensitivity CRP has shown high consistency across a large prospective study underlying the development of unstable atherosclerotic plaques [64]. Other authors also included CRP in the evaluation of incremental usefulness of multiple biomarkers for predicting the risk of cardiovascular events [40, 65].

- **Homocysteine**

  Several studies reported the association between homocysteine levels and different measures of arterial stiffness [64]. However, it was observed that the magnitude of its effect on risk is modest, and the consistency is often lacking, mainly due to nutritional, metabolic (*e.g.*, renal disease) and lifestyle markers [65].

- **Lipoprotein-associated phospholipase 2**

  LpPLA2 has recently emerged as a marker with high associated consistency and precision, and as an independent risk factor in the plaque rupture and atherothrombotic events. Recent studies showed its continuous association with the risk of CHD, with similar magnitudes as the ones verified with non-HDL cholesterol or SBP prevalence [66].

- **Renin-angiotensin-aldosterone system**

  Serum aldosterone and plasma renin are markers of the neuro-hormonal activity and, it is recognised that patients with hyper-aldosteronism present higher arterial stiffness [67]. Despite the role of the RAAS in modulating the arterial compliance, in essential hypertension and in normotensive subjects, is not clear, it has been related to the measurements of the arterial stiffness, such as PWV. It was verified that carotid-femoral PWV is higher in patients with primary aldosteronism, when compared with patients with essential hypertension and with healthy subjects [64].

2.2.3 Prevalence in Portugal

In Portugal, during 2011, 31670 people died ($\approx$ 31 %), from CVD, being 5.9 % reported hypertensive disease, 22 % ischaemic heart disease and, 20 % cerebrovascular disease. These values represent a decreased of 1.6 %, since 2008. The knowledge about the differences in prevalence of the major risk factors by gender and region can influence cardiovascular prevention strategies. Health surveys based on interviews and questionnaires have been the common methodology employed by the Portuguese National Health Survey. Statistical studies have been carried out periodically, since 1987, based on interviews, providing information about risk factors [68].

The AMALIA study [69] was an important epidemiological study carried out in Portugal with the purpose to assess the prevalence and distribution of the main risk factors. The prevalence of the risk factors considered, were identified in the following descending order: sedentarism (76 %), overweight/obesity (51.6 %), hypertension (23.5 %), hypercholesterolemia (19.7 %), smoking (16.3 %), and diabetes (8.9 %). The self-reported prevalence of hypertension was higher in women

(24.9 %) than in men (21.8 %) and, the higher prevalence was verified for obese or sedentary participants, or in individuals with family, or had a family historic of CVD. It is also interesting to note that the smoking prevalence in these three high risk groups was low. However, the same was not true for those with a family history of CVD, probably because of the lower awareness of the risk.

2.3 Diagnostic Tools

There are several methods that can be used for the CVD diagnosis, but only some of them are widely applicable in the clinical settings. A brief review of those methods is presented in this section.

2.3.1 Blood Pressure Monitoring

BP measurement is the basis of the cardiovascular diagnosis, management, treatment, epidemiology and research of hypertension. An accurate BP reading is a pre-requisite, regardless the used technique [70]. The oscillometric method is used for BP measurement, involving the analysis of the oscillations in the cuff pressure. SBP is generally easily identified as the cuff pressure where the oscillations begin. The DBP has not always correspondence with the vanishing of the oscillations and generally the devices employ proprietary algorithms to identify its value [9]. Pereira, H.C. [71] developed an apparatus for pressure values measurement based on accelerometers that is now under provisional patent.

Ambulatory blood pressure monitoring (ABPM) has gradually established its utility, not only in the research field but also in clinical practice, as it allows the study of the BP, Heart Rate (HR) and circadian 24-h BP profile of the patients [72, 73]. The 24-h BP parameters are reproducible and correlate better with target organs damage than the office BP [11]. There is also evidence that ambulatory values predict better cardiovascular morbidity and mortality, rather than office values [74, 75]. Investigations using a photoplethysmographic system have been carried out by several authors, in Massachusetts Institute of Technology (MIT) [76], but also by other research groups [77, 78].

2.3.2 Applanation Tonometry

The applanation tonometry is a non-invasive method that can be applied in superficial arteries that are flattened against a bone structure [79]. This technique is only reliable in peripheral arteries to avoid the effect of the subcutaneous tissues between the tonometer and the artery. The disadvantages of the tonometer include the required operator sensitivity to the sensor placement, because it should be properly immobilized upon the artery of interest, as well as the movement artefacts that may lead to signal distortions, during the data acquisition [79, 80].

2.3.3 Photoplethysmography

Photoplethysmography (PPG) is the more common method used for pulse wave recording, and it is based on the finger blood volume monitoring [9,81]. It uses a light emitter and a detector to measure the pulsatile blood flow. Recent studies have highlighted the use of the PPG signal in blood volume changes monitoring [82], and in pulse waveform analysis. Other studies extended the use of the PPG waveform for the detection of ventricular ejection and for the identification of other relevant pulse transit parameters [83].

2.3.4 Optical Sensors

A simple photoelectric method for measuring vessel diameter changes was developed by Pereira, T. *et al.* [84] by employing distinct types of silicon optical sensors, such as a planar type from Silonex and an avalanche photodiode from Advanced Photonics. Additionally, appropriated light emitting diodes (LEDs) were used as the light source. It was observed that the probes are able to detect and reproduce the waveform distension .

2.3.5 Angiography

Imaging methods can be relevant in CVD risk assessment in individuals with moderate risk. The ”gold standard” imaging method is intracoronary angiography, that allows to localize and to quantify vessels occlusions. This is traditionally done by injecting a radio-opaque contrast agent into the

blood vessel and imaging using X-ray based techniques such as fluoroscopy. However, this method is invasive, carrying risk for the patients [85].

2.3.6 Magnetic Resonance Imaging

Methods based on flow pulse detection using magnetic resonance imaging have been proposed as a non-invasive option for the arterial PWV computation or even local distensibility [18, 86]. However, MRI time resolution is below the one required for flow detection, limiting its clinical application.

2.3.7 Ultrasounds

Risk assessment using carotid ultrasound is based on the measurement of the IMT, and in the presence of plaques and their characteristics [87]. Individuals without known CVD with high increased IMT have high risk of cardiac events and stroke [40].

Arterial diameter can be determined very accurately by ultrasound [88]. Several diameter calibration schemes have previously been reported [89–91]. The ultrasound system measures the intensity of the echo, providing information about the type of structures in the scanned region. The technology is especially accurate in the detection of the interface between solid and fluid filled spaces, which is a key requirement in the vascular analysis. Superficial structures such as vessels, muscles, tendons, breast and the neonatal brain are imaged at higher frequencies (7 to 18 MHz), which provides better axial and lateral resolution. On the other hand, deeper structures such as liver and kidney are imaged at a lower frequencies (1 to 6 MHz), which provides lower axial and lateral resolution, but higher penetration depth [92].

2.4 Parameters used in Arterial Stiffness Assessment

The non-invasive measurement of arterial stiffness implies the measure of surrogate parameters that are intrinsically associated with its development [6, 93]. Three main methodologies are involved: pulse transit time or PWV, analysis of the arterial pressure pulse and its wave contour, and direct stiffness estimation by measuring the diameter and vessel distending.

2.4.1 Pulse Wave Velocity

The PWV measurement is generally accepted as a simple, non-invasive, robust, and reproducible method that can be applied to the arterial stiffness estimation [13]. Measurements along the aortic and the aorto-iliac pathway are the most clinically relevant values. A systematic review, concerning the predictive value of the aortic PWV in cardiovascular events and all-cause mortality, was presented by Vlachopoulos *et al.* [94], where 15877 subjects were included, from the meta-analysis of 17 longitudinal studies. Carotid-femoral PWV was also studied by Laurent *et al.* [95] and by Boutouyrie *et al.* [1]. They proved that the aortic stiffness is an independent predictor of primary CHD in patients with essential hypertension.

There are several algorithms that can be employed on the PWV measurements, being the time delay measured between the foot of two waveforms the most usual [96, 97]. It is also important to consider the heterogeneity of the available algorithms. For example, there is no universal consensus concerning the distance that should be introduced in the PWV equation for its computation [94].

2.4.2 Waveform Parameters

There are several characteristics in the arterial waveform that are indicative of arterial disease [9]. The wave reflections are probably the most studied, often addressed in several studies. The AIx is one of the most relevant parameters in this field, which expresses the ratio of the augmented pressure, assigned to the reflected wave into the overall pulse. Increased AIx and arterial stiffness are highly associated with several cardiovascular risk factors, such as age, smoking, hypertension, diabetes mellitus, and hypercholesterolemia. There is also a strong correlation between them and the high risk of developing CAD.

Another common hemodynamic measurement is the Sub-Endocardial Viability Ratio (SEVR), also known as the Buckberg ratio, that represents the ratio between the myocardial perfusion supply and its demand. It can be calculated measuring the ejection length, determined from the peripheral pulse, the area under the systolic and diastolic portions of the pressure waveform. A lower SEVR may indicate a decrease in the myocardial perfusion.

Further, other parameters can be identified based on the most relevant feature points. Parameters, such as the Reflected Wave Transit Time (RWTT), which is calculated as the elapsed time from the wave foot to the systolic inflection point can be referred as an indirect measurement of the arterial compliance. The Systolic Wave Transit Time (SWTT) can also be computed. Additionally, the systolic ejection time is related with the elapsed time between the beginning of the pulse and the closure of the aortic valve, as indicated by the dicrotic notch, or incisura. This parameters is commonly referred as the Left Ventricular Ejection Time (LVET). Lower ejection duration may be indicative of systolic dysfunction in heart failure patients. A detailed description of these parameters is presented in Section 4.4.6. A review of some studies regarding the association between wave reflections and clinical endpoints is presented in Table 2.1

2.4.3 Diameter Based Parameters

The measurement of the diameter elastic properties of arteries is essential in the study of the dynamic arterial properties. Different techniques have been employed in the distension measurements. Mostly, these techniques are based on electrical, optical and ultrasonic technologies [98, 99]. The arterial distension waveform is almost identical to the pressure waveform [30], but the diameter based waveforms must be converted to pressure, by applying empirical exponential relationships.

Most of these techniques have been used in clinical applications. The simultaneous assessment to the diameter and pressure waveforms have been employed in some studies, allowing the computation of the incremental compliance, distensibility, PWV and elastic modulus [89, 98]. The current commercial devices are essentially used to infer local PWV [100, 101]. However, other authors [90, 102] have developed techniques, based on ultrasound image analysis to assess the systolic local arterial BP, from the analysis of the B-mode diameter waveforms. The evaluation methods require some technical skills and have not been yet generalized to the clinical practice.

Other works, related to the management of neurosurgical disorders, have also considered the waveform morphology of other physiological signals, such as the intra-cranial pressure through the analysis of peak locations [103,104], and the temporal dependency between successive pulses [105].

Table 2.1: Overview of a set of studies on the association between wave reflections and clinical endpoints, sorted by date of publication, where *n* is the number of subjects.

Author		**Clinical Pathology**	**Parameters**	**n**
London *et al.* (2001)	[106]	ESRD	Carotid AIx	180
Ueda *et al.* (2004)	[107]	CAD/angiography	Aortic AIx and aortic Pi	103
Chirinos *et al.* (2005)	[108]	CAD or non-obstructive coronary atherosclerosis	Aortic AIx, augmentation of pressure	297
Weber *et al.* (2005)	[109]	CAD/angiography	HR-corrected aortic AIx (Transfer function)	262
Dart *et al.* (2006)	[110]	Elderly female hypertensives	Carotid AIx	484
Covic *et al.* (2006)	[111]	ESRD	HR-corrected aortic AIx (Transfer function)	92
William *et al.* (2006)	[112]	ASCOT study [a]	Carotid AIx	2199
Manisty *et al.* 2010	[113]	Hypertensive	Carotid AIx, WRI, Pb/Pf	259
Wang *et al.* (2010)	[114]	Normotensive and untreated hypertensive	Carotid measures: AIx, Pi, Pa, RWTT, Pf, Pb	1272
Janner *et al.* (2012)	[115]	General Population	HR-corrected aortic AIx (Transfer function)	3073

ESRD − End-Stage Renal Disease, HF − Heart Failure, Pi −incident pressure wave height, Pa −augmented pressure, RWTT − Reflected Wave Transit Time, Pf− forward wave amplitude, Pb − backward wave amplitude, WRI - Wave Reflection Index

[a] The ASCOT protocol study design has been published in 2001 [116]. Briefly, people were eligible if they were 40 to 79 years of age at randomization and had either untreated hypertension (SBP$\geq$ 160 $mmHg$ or DBP$\geq$ 100 $mmHg$) or treated hypertension (SBP $\geq$ 140 $mmHg$ or DBP $\geq$ 90 $mmHg$) .

2.5 Clinically Approved Devices

The available commercial devices generally estimate the arterial stiffness by the analysis of PWV or AIx values [117].

- **Complior**

  The Complior System (Alam Medical, Les Lilas, France) [118] employs dedicated pressure sensors, directly applied on the skin. It has been widely applied in the carotid-femoral segment [119], and it uses two probes that allow PWV measurement based on the transit time measured between probes. The measurement requires two operators: the first one begins positioning the probe on the common carotid artery, in the central detection site; while the other operator places a second probe at the femoral artery site. The transit time is determined by employing a correlation algorithm between each simultaneous recorded wave.

  Its newer version, the Complior Analysis, offers, in addition to the distension waveform information, the recording of the pressure wave, allowing the assessment of the central pressure and pulse wave analysis [120]. The central BP is not estimated using a transfer function, but using another method based on the second systolic peak (SBP2) of the peripheral waveform. Another relevant difference between both Complior versions is related with the algorithms used for the wave foot determination. In the first version, it was employed a foot-to-foot algorithm based on the wave second derivative, whereas in the last version, an algorithm based on the intersecting tangent is used. The first results from the newer version were presented by Sztrymf *et al.* [120], in a limited population of 12 subjects, and the SBP2 and AIx were only detected in 4 out of the 12 subjects. In spite of the limited number of samples, the results demonstrated that the pressure waveforms are acquired with accurately by this device. The mean difference between the invasive measurements and Complior form factor was 4.2 ± 2.8 % and, differences of 0.4 ± 1.7 % and -0.09 ± 2.8 % were detected for the SBP2 and AIx, respectively.

- **SphygmoCor**

  The SphygmoCor system (ArtCor, Sydney, Australia) [121] is a non-invasive diagnostic device that enables the analysis of the central and ascending aortic pressure wave, using a single high-fidelity applanation tonometer (Millar). The pressure wave is calibrated through a brachial cuff and, it is then converted to the ascending aortic waveform using a generalised mathematical transfer function. The Millar tonometer used by SphygmoCor was introduced in 1989 by Kelly *et al.* [122]. Its performance was assessed by comparing, directly and indirectly, recordings of the radial pressure waves in 62 subjects, of the femoral pressure waves in 3 dogs and of the carotid pressure waves in 17 subjects. It was shown that is possible to monitor the arterial pressure wave contour under a wide variety of clinical conditions and pulse pressures. However, small differences were observed in the frequency analysis, namely, the trend to overestimate the 1st harmonic and to underestimate moduli above 4th harmonic [120]. SphygmoCor is also capable to perform measurements of a wide number of parameters, such as AIx, PWV, Ejection Duration (ED), HR, dp/dt, SBP, DBP, SEVR.

  Several authors have validated SphygmoCor system measurements, by comparing the peripheral to central aortic transfer function data, against transducer-tipped catheter measurements [123, 124], while others have been dedicated to the study of the calibration algorithm [125]. The accuracy of the aortic AIx determination showed a consistence, presenting a moderated underestimation of the aortic AIx [126]. In what concerns the pressure values, Cloud *et al.* [125], obtained on average, an underestimation of the SBP in 13.3 mmHg and an overestimation of the DBP in 11.5 mmHg.

- **Finapress**

  Finapress (Finapres Medical Systems, Amsterdam, The Netherlands) uses the volume-clamp method, in which the finger artery is gauged with an infrared transmission plethysmograph mounted inside an inflatable cuff. When the size of the artery increases due to the higher blood volume (and pressure), the air pressure in the cuff also increases, in order to keep the arterial size constant. The arterial size and the transmural pressure are kept constant, by modulating the cuff pressure in parallel with the BP in the finger, through a pneumatic servo system.

- **Arteriograph**

  The Arteriograph (TensioMed Kft., Budapest, Hungary) is an oscillometric based device used for BP, AIx and PWV computation [117]. The oscillometric pressure curves are registered in the upper arm (brachial artery). Initially, the BP is measured in the upper arm, followed by the increase of the cuff pressure, 35 mmHg above the SBP measured. For the analysis of the pulse wave, the time difference between the beginning of the first wave and the beginning of the second (reflected wave), required for the PWV measurement, is identified. Then, the AIx is computed by the amplitude difference between the first and second wave in relation to the pulse pressure.

  The PWV and AIx values, measured using the Arteriograph, were compared with the corresponding Complior and SphygmoCor measurements, in 254 untreated hypertensive patients. It was verified that the Arteriograph PWV and AIx values are closely related, when compared with Complior ($r = 0.60,\ p < 0.001$) and SphygmoCor ($r = 0.89,\ p < 0.001$).

- **PulsePen**

  The PulsePen (Diatecne, Milan, Italy) is an instrument used in several population studies and clinical trials, for the assessment and analysis of the pressure waveforms, and for the measurement of arterial distensibility [127, 128]. It is composed by one tonometer and an integrated electrocardiogram (ECG) unit, and it provides the PP values. The system is calibrated to the mean arterial pressure value and the DBP is obtained using a conventional method. Salvi, *et al.* [129] performed a clinical validation, by monitoring the pulse waves in the carotid artery, using the PulsePen *versus* the intra-arterial simultaneous measurements, in 10 patient undergoing cardiac catheterization. The invasive signals were recorded at the origin of the ascending aorta, and at the initial tract of the carotid artery, using fluid-filled catheters; while non-invasive waveforms were recorded at the common carotid artery and at the base of the neck (bottom carotid). The comparison between both methods showed no differences in spectral moduli from harmonics 1 to 6. Additionally, a good PWV performance was verified from the data collected with the PulsePen, in comparison with the Millar tonometer measurements in 68 subjects

2.5.1 Standardization of Measurements

The standards for evaluation medical measurement devices are set by three main professional societies: the Association for the Advancement of Medical Instrumentation (AAMI), the British Hypertension Society (BHS) and the European Society of Hypertension (ESH) [9]. The oldest standards for devices validation are related to the sphygmomanometers and were published by AAMI in 1986.

In what concerns the waveform analysis, the existence of devices using different technologies, renders the use of waveform based parameters difficult in clinical practice, namely in the elaboration of proper reference values. Comparative studies have been made by several authors that evaluated the correlation of data obtained with different devices. Salvi *et al.* [130] compared the Complior and the PulsePen devices and, verified that the PWV is significantly lower when measured by Complior. These differences occur due to the use of different algorithms in the computation of the foot-to-foot time delay. This study also shows that is not possible the definition an absolute reference for PWV, without consider reference values for each one of the available devices. In another study from Jatoi *et al.* [117], it was also verified a poor agreement between the PWV and AIx measurements for devices (Arteriograph, Complior and SphygmoCor).

2.5.2 Accuracy Validation Studies

Beside the biological variability of a human being, the reproducibility depends on two technical factors. On the one hand the variation of the measurement method and, on the other the effect of the operator. Several investigations can be found regarding this topic for several devices in literature.

In order to apply reproducible data in clinical practice, three levels of acceptance of the measurements have been used in literature. Good variability is defined below 10 %, variability values between 10 % and 30 % are considered satisfactory, but dependent of the use in practice, the cost, and the availability of alternative methods, and finally, variabilities over 30 % are considered unsatisfactory [131, 132].

Unfortunately, the comparison between measurements from different devices is hard, since different statistical measures are used by each one of the studies, such as the Coefficient of

Variation (CV), Standard Deviation (SD), Repeatability Coefficient (RC) or 95 % Confidence Interval (CI). The intra-observer variability can be assessed by the same observer on different but close occasions and the inter-observer variability correspond to the measurements performed by independent observers. The day-to day variability is another typical measurement that can be assessed by the same observer at different occasions (after few days or weeks). A comparison of obtained values from relevant published literature is presented in Table 2.2.

Table 2.2: Comparison of variability values obtained from different commercial devices.

Device	Parameter	Variability	Test	Values	n	Ref
Complior	PWV	Intra-observer	CI	3.30 %	20	[131]
		Inter-observer		2.60 %		
		Day-to day		5.60 %		
	PWV	Intra-observer	RC	0.935	56	[119]
		Inter-observer		0.89		
PulsePen	PWV	Inter-observer	CV	7.94 %	68	[129]
		Intra-observer		7.20 %		
	AIx	Inter-observer	CV	15.80 %		
		Intra-observer		15.15 %		
SphygmoCor	PWV aor	Inter-observer	Bland-Altman ($mean \pm SD$)	$0.3 \pm 3.2\ m/s$	19	[133]
		Day-to-day		$-0.7 \pm 1.9\ m/s$		
		Intra-observer1		$-1.0 \pm 3.9\ m/s$		
		Intra-observer2		$-0.1 \pm 3.6\ m/s$		
		Intra-observer3		$-0.3 \pm 2.4\ m/s$		
	AIx	Inter-observer		0.9 ± 15.8 %		
		Day-to-day		2.6 ± 11.2 %		
		Intra-observer1		2.7 ± 18.8 %		
		Intra-observer2		1.9 ± 10.6 %		
		Intra-observer3		0.9 ± 9.3 %		
	RWTT	Inter-observer		$-1.9 \pm 30.8\ ms$		
		Day-to-day		$-2.4 \pm 12.3\ ms$		
		Intra-observer1		$-6.9 \pm 52.7\ ms$		
		Intra-observer2		$-2.7 \pm 32.8\ ms$		
		Intra-observer3		$-3.2 \pm 33.9\ ms$		
	AIx	Inter-observer	Bland-Altman ($mean \pm SD$)	0.4 ± 6.4 %	25	[134]
		Intra-observer1	($mean \pm SD$)	14.8 ± 12.7 %		
		Intra-observer2		12.6 ± 14.3 %		
		Intra-observer3		12.7 ± 14.0 %		
		Intra-observer4		12.8 ± 14.1 %		
		Intra-observer5		9.4 ± 14.0 %		
		Intra-observer6		12.2 ± 12.0 %		
	AIx@75HR	Inter-observer	Bland-Altman ($mean \pm SD$)	1.0 ± 3.9 %	20	[135]
		Intra-observer		-1.5 ± 7.0 %		
		Intra-observer		0.1 ± 8.0 %		
	AIx	Intra-observer	Bland-Altman (SD)	7.50 %	302	[136]
Ultrasound (Aloka Pround)	AIx	Intra-observer	CI	17.80 %	20	[131]
		Inter-observer		8.40 %		
		Day-to day		23.80 %		

CV−Coefficient of Variation, SD−Standard Deviation, RC−Repeatability Coefficient, CI−95 % confidence interval

2.6 Reference Values

The purpose of this section is to provide reference values from a survey of the literature about normal and reference values for the set of parameters used in this document, namely information about the BP values, the biochemical markers levels, as well as, indications about the PWV and APW parameters.

2.6.1 Blood Pressure

Elevated BP is a major risk factor for several diseases, such as heart failure, cerebrovascular disease, renal failure and atrial fibrillation. The standard definition and classification of the hypertension levels is shown in Table 2.3.

Table 2.3: Definition and classification of hypertension. Values based on [40].

Category	**SBP (mmHg)**	**Condition**	**DBP (mmHg)**
Optimal	< 120	and	< 80
Normal	$120 - 129$	and/or	$80 - 84$
High Normal	$130 - 139$	and/or	$85 - 89$
Grade I hypertension	$140 - 159$	and/or	$90 - 99$
Grade II hypertension	$160 - 179$	and/or	$100 - 109$
Grade III hypertension	≥ 180	and/or	≥ 110
Isolated systolic hypertension	≥ 140	and	< 90

2.6.2 Body Mass Index in Adults

The Body Mass Index (BMI) has been extensively used in the body weight analysis. In adults, overweight is defined by a BMI ranging from 25 to 29.9 kg/m^2, and obesity by a BMI ≥ 30 kg/m^2. It is proved that high BMI is associated with risk of CVD [137].

2.6.3 Dyslipidemia

Dyslipidemia is connected with the abnormal amount of lipids in the blood. The most common alterations are related to the increase of the plasma cholesterol and LDL cholesterol levels. Hypertriglyceridemia and low HDL cholesterol levels are also independent CVD risk factors. The threshold risk values are presented below:

- **Plasma cholesterol** - In general plasma cholesterol should be $< 190\ mg/dL$ [40].
- **LDL cholesterol** - LDL cholesterol should be $< 115\ mg/dL$ [40].
- **HDL cholesterol** - HDL cholesterol $< 40\ mg/dL$ in men and $< 45\ mg/dL$ in women may be regarded as a marker of high risk [40].
- **Tryglicerides** - TGL value $> 1.7\ mmol/L$ is the threshold to be considered a marker of high risk [40].

2.6.4 Augmentation Index

The peripheral and central AIx values increase with age and with the mean arterial pressure, and are inversely related to the HR and body height. Janner *et al.* [138] studied 4561 subjects in *The Copenhagen City Heart Study*, and calculated reference values for the AIx, measured with the SphygmoCor. AIx reference equations considering age, HR and height were defined, for both men (Equation 2.1) and women (Equation 2.2):

$$AIx_{men} = 79.20 + 0.63(age) - 0.002(age^2) - 0.28(HR) - 0.39(height) \tag{2.1}$$

$$AIx_{women} = 56.28 + 0.90(age) - 0.005(age^2) - 0.24(HR) - 0.34(height) \tag{2.2}$$

Recently, many studies have AIx values in different races. Table 2.4 present some results obtained from the analysis of different populations. Proposed values in South African [143], Chinese [141] and Korean [142] populations are higher, comparatively to those estimated in Europeans. This discrepancy might be due to the differences in the mean age of the populations (the European group

Table 2.4: AIx reference values for different populations

Population	Gender	Mean Age (y)	AIx (%)	n	Ref
European	Women	35.1 ± 13.6	15.1 ± 16.3	306	[139]
	Men	34.6 ± 13.4	7.2 ± 14.6	228	
American	Women	57 ± 8	17 ± 10	333	[140]
	Men	56 ± 9	9 ± 12	188	
Chinese	Women	38.7 ± 11.6	21.0 ± 14.4	456	[141]
	Men	42.7 ± 13.2	16.6 ± 15.6	468	
Korean	Women/Men	36.3 ± 9.6	28.2 ± 43.0	522	[142]
African	Women/Men	33.5 ± 13.1	22.2 ± 14.3	185	[143]

was younger), as well as due to the ethnic differences (*e.g.*, the shorter stature evident of Korean subjects). Wojciechowska *et al.* [139] (2006) observed central AIx values, for the European subjects, very similar to the values obtained by Mitchell *et al.* (2004) [140], in the Framingham Heart Study Offspring Cohort, despite the predominance of older subjects in the first study.

The detailed threshold values for different ages, presented in the European study [139], can be seen in Table 2.5. The central AIx values were on average lower among men subjects. The authors also observed that the values were curvilinearly related to age, but similar in men and women.

Table 2.5: AIx reference values for healthy white European subjects, measured using SphygmoCor.

Gender	Age Group	n	AIx (%)
Men	< 30	106	-1.5 ± 12.1
	30-39	51	9.4 ± 10.7
	40-49	33	13.7 ± 11.4
	≥ 50	38	23.1 ± 10.6
Women	≤ 30	134	3.7 ± 14.3
	30-39	64	16.6 ± 10.2
	40-49	56	25.6 ± 10.6
	≥ 50	52	31.1 ± 8.7

Comparisons between AIx and PWV measurements (considered the "gold standard" parameters for arterial stiffness estimation), have also been reported. Data from a large cohort of healthy individuals, in the Anglo-Cardiff Collaborative Trial (ACCT), showed that the central AIx might be a more sensitive marker of arterial ageing in young and middle-age individuals ($<$ 50 years), and that the aortic PWV is more sensitive in older individuals ($>$50 years) [144].

2.6.5 Pulse Wave Velocity

Commonly, the norms and reference group data, used for the statistical analysis are established according to the health status, age and by combined hypertensive-diabetic status, but more rarely to the hypertensive and normotensive degree. In Table 2.6 the results achieved by Elias, M. *et al.* [145] are presented, from a sample $(n = 502)$ ranged from 40 to 93 years old, after exclusion of individuals with acute stroke, probable dementia, and diabetes history. The rise of the PWV values with age and hypertension was observed, as expected.

Table 2.6: PWV reference values obtained using SphygmoCor system [145].

Type	**Age group**	**n**	**Value (m/s)**
	40-49	34	8.1 ± 1.3
	50-59	69	8.0 ± 1.3
Normotensive	60-69	48	8.9 ± 1.9
	70-79	40	10.1 ± 2.1
	80-83	15	11.4 ± 2.6
	40-49	17	8.6 ± 1.5
	50-59	63	9.3 ± 1.5
Hypertensive	60-69	85	10.3 ± 2.3
	70-79	84	12.3 ± 3.1
	80-83	47	13.1 ± 3.1

2.6.6 Waveform Analysis

Few studies regarding the waveform reference values can be found in the literature, due to the novelty of its concept in the scientific community and the inherent difficulty in gather a population size statistically significant. The studies in this area are mainly dedicated to RWTT, SWTT and LVET analysis. A brief review of the observed values, organized by population groups, is summarized in Table 2.7.

Table 2.7: Timing reference values for the RWTT, SWTT and LVET parameters.

Parameter	**Reference**	**Gender**	**n**	**Value**	**Ref**
RWTT (ms)	Frimodt-Moller M. *et al.*	Men/Women	19	137.1 ± 32.1	[133]
	Mitchell G. *et al.*	Men	188	138 ± 28	[140]
		Women	333	124 ± 28	
SWTT (ms)	Mitchell G. *et al.*	Men	188	201 ± 53	[140]
		Women	333	210 ± 29	
LVET (ms)	Mitchell G. *et al.*	Men	188	312 ± 23	[140]
		Women	333	320 ± 22	

2.7 Risk Assessment

The international guidelines [40, 146] consider that individuals with established CVD should be the first priority for preventive measures application. The concern in changing the current healthcare paradigm, from reactive towards preventive care, aims at identify individuals for risk in early stages, and then, direct more efforts and attention to the risk factors modification [147, 148]. Fortunately, this is an emergent tendency that can be addressed using the traditional risk scores or machine learning techniques.

2.7.1 Risk Scores

During the last years many risk estimation systems have been developed in order to assist clinicians in the risk factor assessment, and in the individual chances prediction, for CVD development. The major challenges of these tools are: (1) identify high risk individuals, (2) weight the individual effects of all risk factors, (3) stratify or organize who needs lifestyle advice or medical therapy, and finally (4) avoid overmedicalization of individuals at low risk [149]. Taking this challenges into account, several risk factors were identified, by their association with an increased risk of CVD development, such as atherosclerosis.

CVD risk assessment tools differ from each other on the selected risk factors, disease for what they were designed (coronary heart disease, heart failure, *etc.*), selected event type, considered period of time (long or short term) for the prediction and cohort location. The most popular are: FRS, PROCAM, ASSIGN and SCORE.

2.7.2 Machine Learning Techniques

Computer-aided diagnosis methodologies, based on data mining and machine learning techniques, have been used in many biomedical applications during the last years [150–153]. Amongst the many different definitions that are found in the literature, data mining can be described as the automatic or semi-automatic process for discovering patterns in huge amounts of data. These techniques could contribute to the development of clinical decision-making approaches, using information taken from different measurement setups, overcoming the limitations in the management of complex and unstructured data [17, 150, 154, 155].

2.7.2.1 Classification Techniques

An interesting approach is the exploration of different classifiers, as it was proposed by Jovic *et al.* [151]. The ECG classification problem was addressed using a combination of several features in the analysis of the HR variability. Other approach presented by Tsipouras *et al.* [150] was based on the development of a fuzzy rule-based decision support system for CAD diagnosis. On the

other hand, multi-classifiers should perform better in some situations, overcoming errors from single classifier analysis [156]. The incorporating of the prediction outcome of each one of the individual classifier was suggested [157], as a way to reduce the classification errors.

2.7.2.2 Clustering techniques

Clustering analysis is another important branch of unsupervised learning that allows the arrangement of objects into groups (*i.e.*, the clusters), wherein the objects in the same cluster are more similar (in one or more characteristics), than those in different clusters [158]. There is a wide variety of clustering methodologies available in literature, essentially organized in three general classes [159]. The three types include parametric model-based, hierarchical and partitioning algorithms.

Shah, *et al.* [160] has proved the usefulness and feasibility of using clustering risk factors in the detection of CVD in youth, by the comparison with the Patholobiological Determinants of Atherosclerosis in Youth (PDAY) risk score. Other studies have also referred the role of clustering methodologies in CVD assessment, such as the work developed by Haseena *et al.* [161], where a fuzzy C-mean clustered probabilistic neural network for ECG beats discrimination was described. Clustering methodologies were also applied in other medical fields, such as in the identification of patterns in blood glucose measurements and regular insulin doses taken before meal time [162].

3

Multichannel Biomedical Platform

In this chapter the multichannel biomedical platform is presented. The platform is based on several modules interfaced to a standard PC for data processing, storage and display. A brief discussion about the power management circuitry used to ensure the required autonomy of the portable instrument, as well as the database connectivity and synchronisation mechanism between all of the modules is presented. The APW module is described in detail, focusing on the hardware blocks, the implemented firmware and the mechanical design. The system performance was assessed in a test bench.

3.1 Motivation

The underlying idea for the development of a multi-modular integrated instrument is the possibility to sum up several technologies and methodologies used in the traditional clinical path of cardiovascular patients in order to provide a refined assessment of their physiological status. Such accomplishment is granted by the combination of traditionally available clinical information (ECG, PPG), usually not fully correlated, with the one obtained with newly developed non-invasive instrumentation modules. Together, these modules can provide a detailed profile of the arterial pulse structure.

Several modules were developed in Electronics and Instrumentation Group - GEI (Instrumentation Centre of the University of the Coimbra - Portugal), and in Institute of Nanostructures, Nanomodelling and Nanofabrication - I3N (University of Aveiro). These modules are dedicated to PWV assessment (double piezo and sound probes) [96, 163], ECG monitoring, PPG measurements [164], and pulse waveform measurements, using optical [84, 165], piezoelectric [166] and fiber-Bragg sensors [167]. Some of these modules have been implemented in the multichannel platform, shown in Figure 3.1.

Figure 3.1: Multichannel biomedical platform.

3.2 System Architecture

This multichannel platform was implemented using mainstream technologies and protocols, mechanically based on a Schroff crate (24572-003 - 3U, 315MM, 42HP). A schematic view of the system is illustrated in Figure 3.2, where it is possible to identify the main components. The electronics rack provides the core of the real-time hemodynamic system, using a custom made backplane bus for communication between the individual modules, and grants the interface with the computer, and with the autonomous power supply (PS unit), a rechargeable 12 V lead acid battery. A detailed description of each one of the developed modules is presented in the following sections.

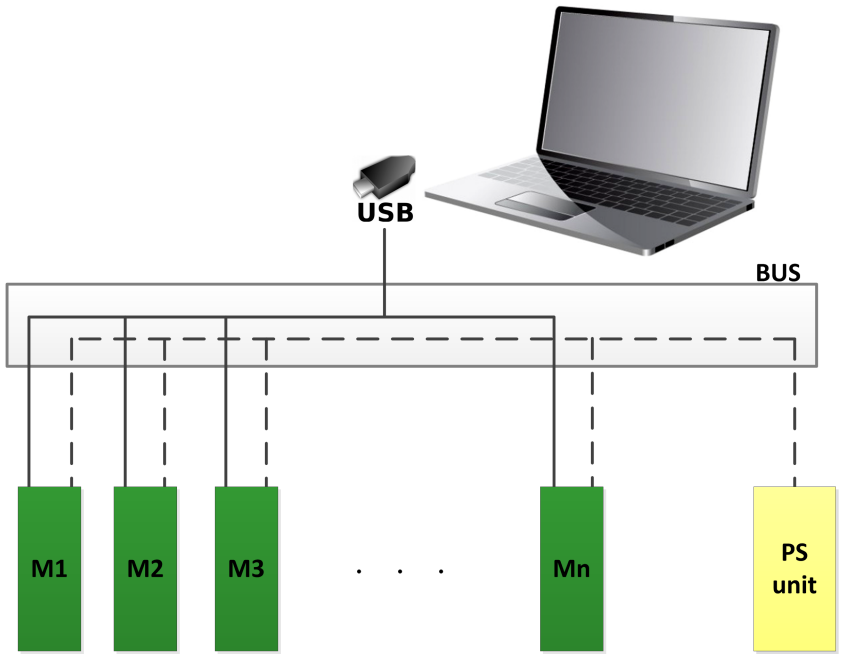

Figure 3.2: System architecture of the multichannel platform, where M1, M2, M2, Mn are the individual models, and PS is the power supply unit.

3.2.1 Power Supply Module

The entire electronics box is designed to run from a single 12 V DC power source, using an embedded lead acid battery. From this power source several other DC power supplies lines are derived using DC-DC converters. These supply lines are $+5\ V$, $+15\ V$, $-5\ V$, $-15\ V$ and $+3.3\ V$. Each of these supplies is routed along the backplane to feed all of the modules. It is also ensured common ground potential lines.

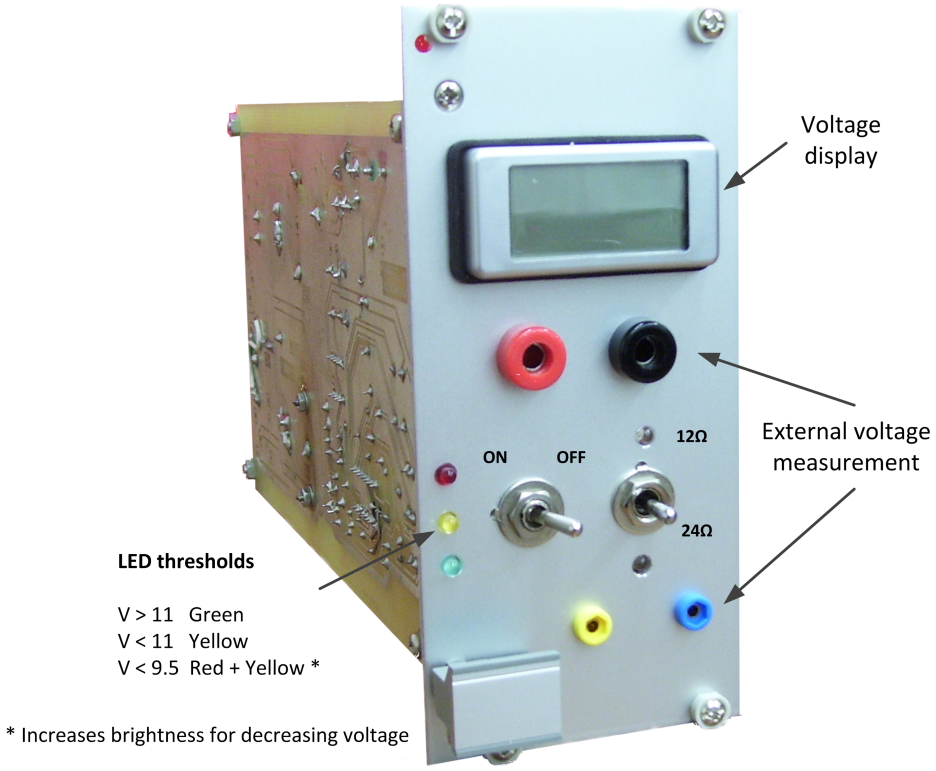

Figure 3.3: Battery power test module.

Two power module versions were implemented. The first one with visual alarms that are activated when specific voltage thresholds are reached (see Figure 3.1 - module 1). And another, only available for laboratory trials, which allows the battery status monitoring during the discharge process (Figure 3.3).

3.2.1.1 Autonomy

Tests on the battery autonomy were performed for two different discharge loads. It was considered the discharge time, expressed as hours elapsed until a fully charged cell reaches the cut voltage and has to be recharged. Figure 3.4 shows the constant current discharge behaviour of the battery with $V_{begin} = 12.5\ V$ and $V_{cut} = 11\ V$. The blue and black curves in the plot correspond to the discharge currents for loads of 24 Ω and 12 Ω, respectively. When the battery was subjected to a continuous discharge current of $\approx 1\ A$, it reached an autonomy of about 3 hours, while when subjected to a discharge current of $\approx 500\ mA$, it reached an autonomy of nearly eight hours.

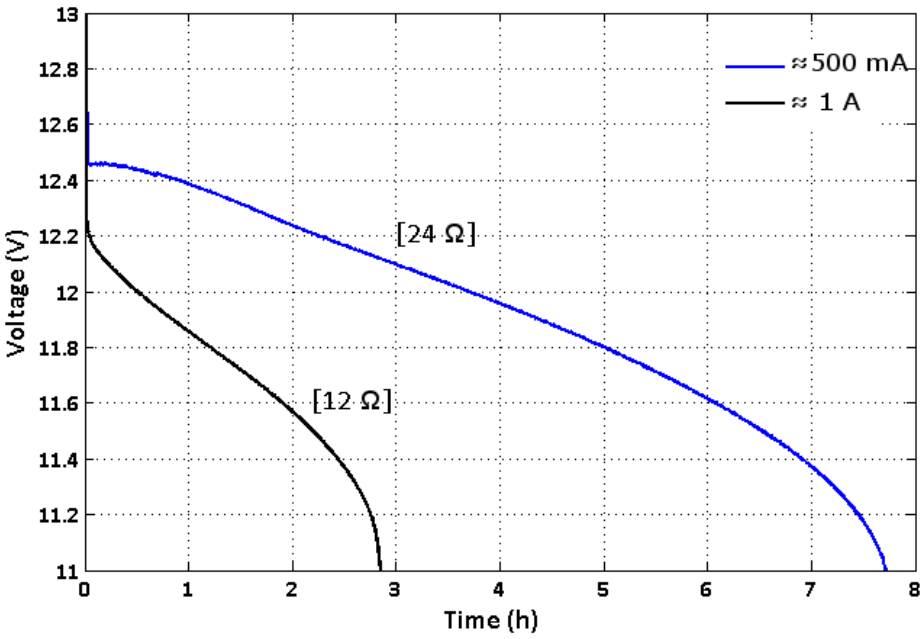

Figure 3.4: Discharge behaviour of the load acid battery with $V_{begin} = 12.5\ V$ and $V_{cut} = 11\ V$.

3.2.2 Bus

The backplane (100x100 mm^2) printed circuit board (PCB) is the only physical communication channel between the various cards inserted in the rack. Connections to each card are provided via six 64-pin DIN bus connectors. Five of these connectors are spaced at 4HP (1HP= 5.08 mm) from each other, while another is at a distance of 6HP. It is an entirely passive device with three main groups of signals. Fourteen lines are used for DC power delivering to each one of the cards, as outlined in Section 3.2.1. The remain 52 pins of each connector are connected to the parallel bus, used in the communication between the cards: 16 lines to analogue communication, and the remaining are dedicated to digital channels. Its configuration is schematically represented in Figure 3.5.

3.2.3 Individual Models

The APW module was initially considered to operate as a standalone module. However, the clinical trials evolved to a point, where it was considered the development of a multi-parametric integrated instrument able to include other technologies useful for the characterization of the patients.

The implemented modules include: the APW module (presented in detail along this document),

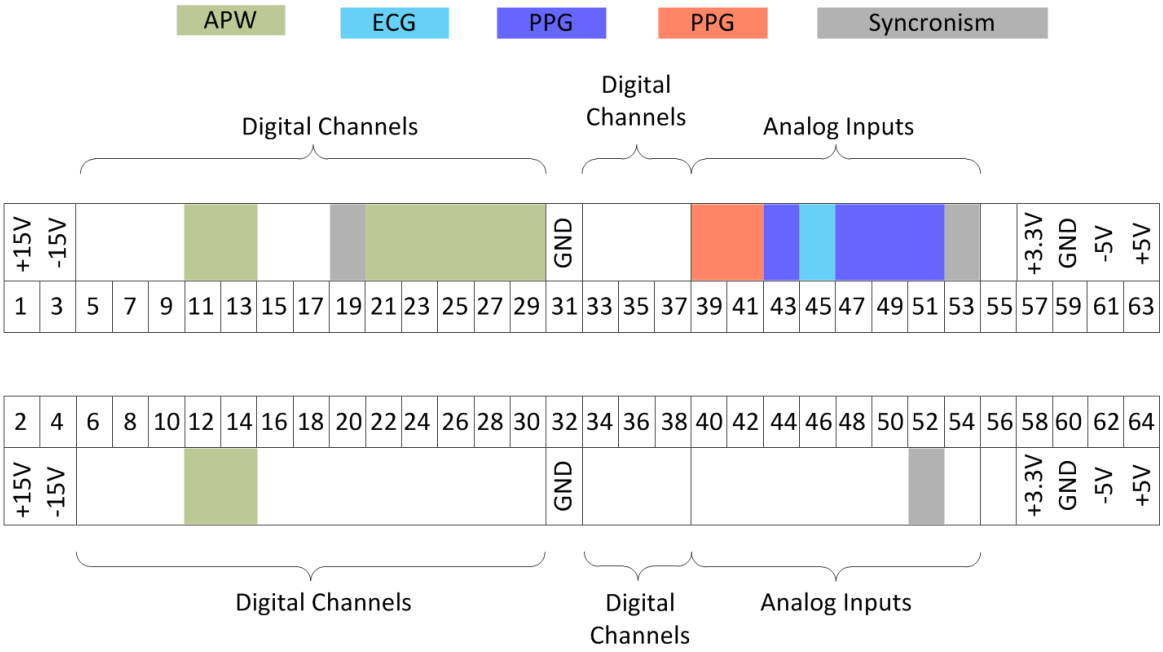

Figure 3.5: Backplane bus configuration.

the PPG module (developed by Pedro Santos [164]), the PWV module (developed by Catarina Pereira [96]) and the ECG module (developed by Prof. Dr. Requicha Ferreira). Table 3.1 presents the main characteristics associated to each of the modules. The PPG, PWV, ECG and NI 6210 (used as the data acquisition board) modules are shown in Figure 3.6. The APW module is described in detail in Section 3.3.

Table 3.1: Modules used in the multichannel platform.

Module	**Description**	**Sample rate**$_{(min)}$
NI 6210	Multifunctional DAQ module that offers 16 analogue inputs, a 250 kS/s single-channel sampling rate, 4 digital input lines and 4 digital output lines	-
PWV	Double-PZ probe devoted to PWV measurements	20 kHz
ECG	3-Lead ECG module with a 1st-order notch filter (50 Hz)	1 kHz
PPG	Photoplethysmographic module based on red and infrared (IR) LEDs	1 Hz
APW	Dedicated system to APW (single PZ probe), and respiration monitoring. It also offers an extra channel for a double PZ probe	1 kHz

Figure 3.6: Detailed front panels of some of the modules: 1-PPG module, 2-ECG module, 3-PWV module, 4-NI 6210 module.

3.2.4 Synchronization

Synchronization is a key factor in the performance of the multichannel platform, since there are signals acquired by different DAQ components (the NI6210 from National Instruments® and the PIC33FJ256GP710 from Microcip®). This means that signals might not be properly synchronized, in time, due to offsets or differences in the different acquisition rates, unless a proper sync procedure is performed.

A synchronization algorithm was developed by Vaz, P. [168], based on the cross-correlation between signals of the different modules. This method is based on the random nature of white noise and, its large frequency band, allowing the determination of the delay between modules and, the establishment of a common temporal reference.

3.2.5 Database

All the retrieved clinical parameters obtained, from each patient with one or more modules, are concentrated on a secure [1] dedicated relational database, developed by Santos, P. [169], along with their corresponding historical, physiological and biochemical data. The database was implemented using MySQL.

Cardiocheck main menu has a set of shortcuts that can be used to select common general tasks such as acquisition, search, creation of instances, data processing and settings (Figure 3.7). Other relevant informations are presented in Appendix A.

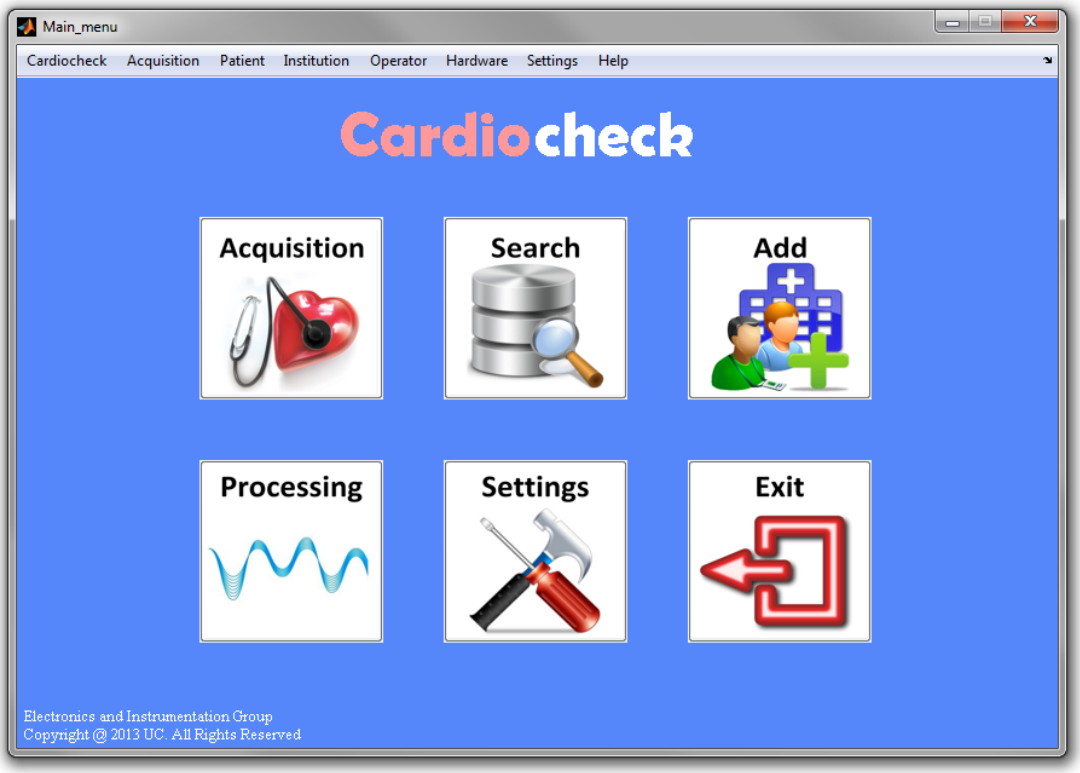

Figure 3.7: Cardiocheck main menu.

[1] These data are strictly confidential.

3.3 Arterial Pressure Waveform Module

Three architectural designs have been experimentally characterized through prototype versions with timely incrementing features (each re-designed taking advantage of the previously achieved experience). The final prototype (DAQFP) was implemented in the multichannel platform. The main characteristics of each one of the prototype versions are summarized in Table 3.2, and the corresponding pictures of the first two versions are shown in Figure 3.8.

- **First Prototype Version**

  The First Prototype Version (DAQP1) system was a first design attempt, in order to assess the performance of PZ sensors in APW monitoring. The NI 6210 DAQ board was used. This first prototype configuration contributed for the selection of the best hardware topology, namely: electronic signal conditioning test circuits, front-end electronics, PZ signal recovery amplification circuit, and baseline restorer (BLR) circuit (see 3.3.1.5). Several limitations were found during the test of this version in subjects, mainly the difficulty to deal with the physiological variability between subjects.

- **Second Prototype Version**

  The Second Prototype Version (DAQP2) was the first proof of concept prototype capable to perform the pulse-by-pulse APW data acquisition. This prototype contributed to debug and customize DAQ firmware, as well as, to settle the basis for the acquisition software. It is based on a custom designed acquisition and processing board from the Microchip®.

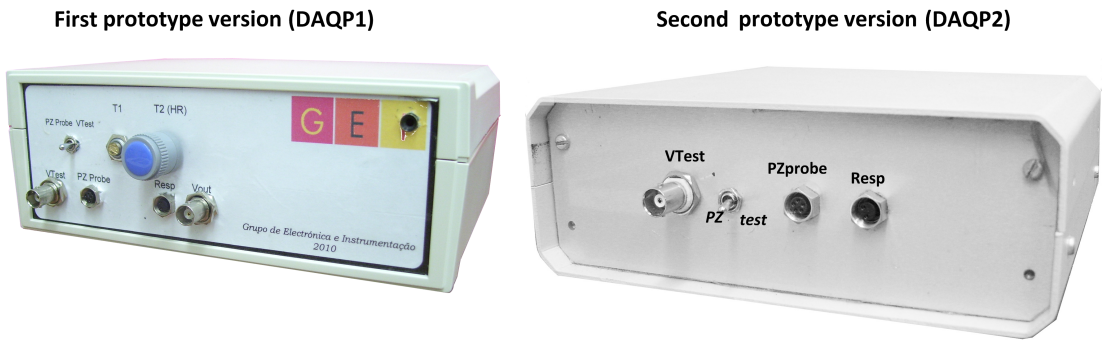

Figure 3.8: First and second prototype versions developed along this thesis.

- **Final Prototype**

  The DAQFP is the final version of the laboratory prototype in which all of the two previous versions converged. This module was also designed to be integrated in the multichannel platform. The board assembly was performed in a 100x60 mm^2 standard Eurocard board, with 64 pins DIN plug and, with an appropriate front panel plate. Additionally, this version is powered by an external autonomous power supply. A detailed view of the designed PCB of this module is presented in Figure 3.9

3.3.1 Hardware Architecture

The PZ sensors were the first block to be explored, due to their fundamental role of being the signal source interface between the patient and the acquisition module. In the next sub-section, a brief description about PZ sensors is presented and, then, the full discussion of each building block of the

Table 3.2: Main characteristics of DAQ prototypes.

Prototype version	**Date**	**PCB Area**	**Characteristics**
1st prototype (DAQP1)	April 2010	100 cm^2	– Manual reset – Baseline restoration based on an analogue circuit – NI6210 board – USB supply
2nd prototype (DAQP2)	Nov 2010	152 cm^2	– Automatic reset – Baseline restoration based on a digital process – PIC board – USB supply
Final prototype (DAQFP)	July 2012	160 cm^2	– Modular platform – Connectivity with other modules – Synchronization – Battery supply

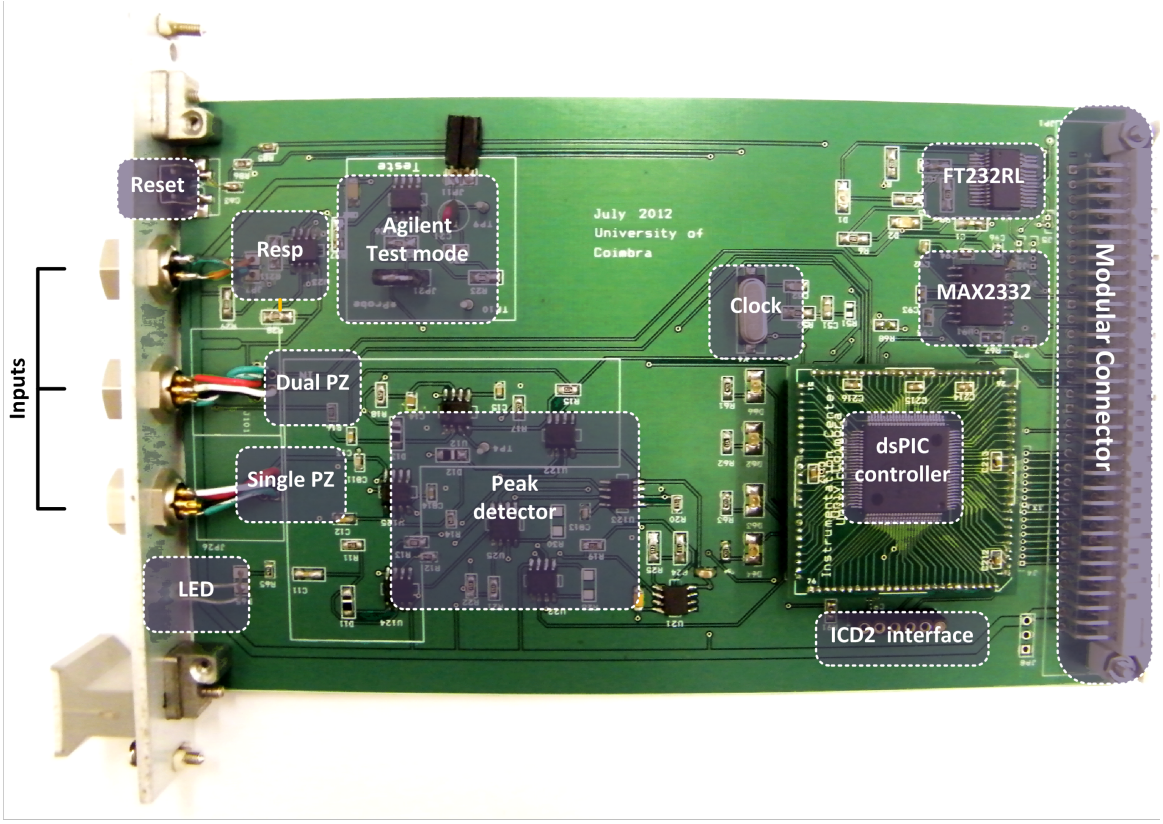

Figure 3.9: The final version of the APW module prototype based on the *dsPIC*-based core, housed in a eurocard (100x160 mm^2) format.

complete electronic circuit enclosed in the shaded boxes of Figure 3.10 is presented: the PZ sensors, the first amplifying stage that connects to the sensor; the peak detector and timer; the integrator with reset; and the BLR (responsible for eliminating very low frequency fluctuations of the foot of each pulse). Additionally, the breathing circuit and the *dsPIC* circuit implementation, that are also part of the APW module, such as represented in Figure 3.9, are presented.

3.3.1.1 Piezoelectric Sensors

PZ sensors exhibit the capability of delivering very high Signal-to-Noise Ratio (SNR) voltage signals when small forces are applied. Figure 3.11 shows the *Thevénin* equivalent lumped-parameters of the PZ converter (a) and its simulated frequency response (b). Due to its construction process, the electric equivalent of a PZ sensor is, essentially, a capacitive impedance (C_S) in series with additional elements that account to the wiring inductance and resistances (Lm, $Rm1$ and $Rm2$) and an eventually associated electrical load (Rk and Ck). In this circuit: VS is the voltage replica to the external

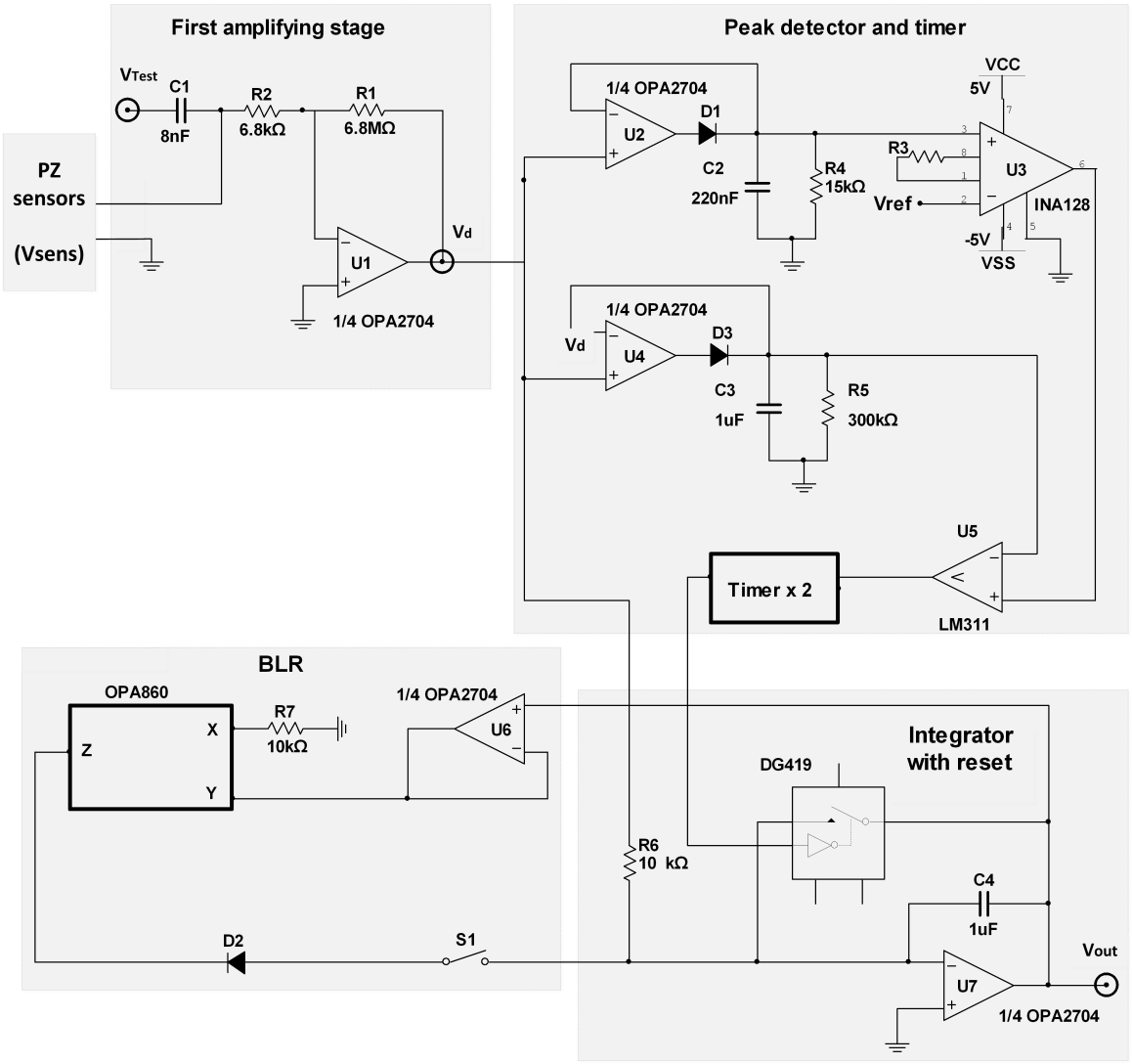

Figure 3.10: Complete circuit with its main functional blocks enclosed in the shaded boxes. V_{Test} and V_{Out} are the voltages from which performance of the circuit is evaluated. V_d is a monitoring point for the circuits studied in 3.3.1.2. These voltages are outlined by circles.

mechanical excitation, $x(t)$; R_S is the insulation (or leakage resistance) generally in the 10 to 50 $M\Omega$ range, and C_S is the series capacitance which, among others, depends on the characteristics of the PZ element and geometry of the sensor. For the device used in this work this value stays around 8 nF [171]. This value is consistent with the capacitance of a disc capacitor of 1 cm of radius and with the distance between plates of 0.5 mm, and a 2000 relative dielectric constant.

If the frequencies of interest are confined to the flat *plateau* of the frequency response (area limited to the dashed box in Figure 3.11(b)) the circuit can be satisfactory modelled as the simple series capacitor, C_S. This is the case in hemodynamic studies, where the bandwidth of signals is well

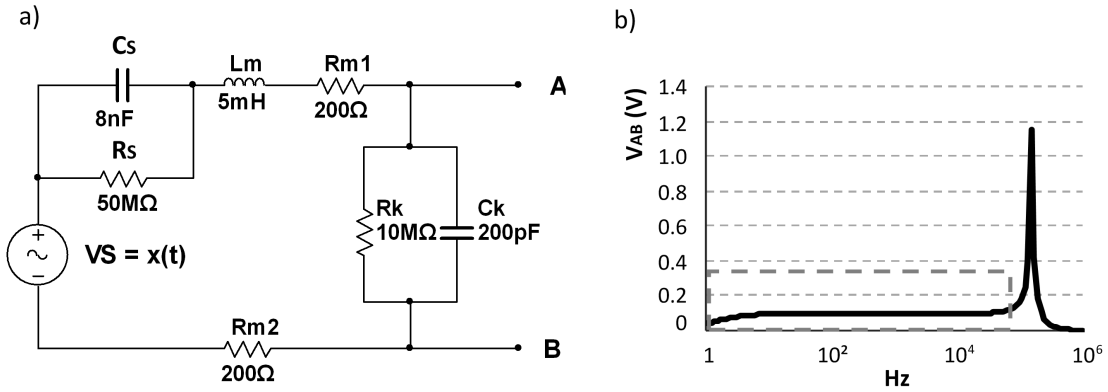

Figure 3.11: PZ sensor characteristics: (a) *Thevénin* equivalent based on [170] and (b) its frequency response, obtained by simulation.

below the frequency of the resonance peak ($4.6 \pm 0.5\ kHz$ for the Murata® 7BB-27-4 in use [171]). This model also explains the fact that PZ sensors act as AC coupled voltage generators (note that R_S is the modelled insulation resistance, similar to the one specified in most capacitor datasheets) and, consequently, do not respond to static excitation. It also gives a lead for two possible amplifying topologies for the electronic circuit that immediately follows the sensor: the Charge Mode Amplifier (CMA) and the Voltage Mode Amplifier (VMA) [170].

3.3.1.2 Amplifying Stage

This block accepts inputs from two different sources: one originated from the sensor, V_{Sens}, and another one, usually derived from a signal generator, V_{Test}, used for test purposes (see Figure 3.10). As the aim is the precise rendering of the pressure waveform, $P(t)$, this section is devoted to the identification of three topologies to achieve this goal. Firstly, due to the inherently capacitive nature of the sensor, depicted in the equivalent circuit of Figure 3.11(a), its voltage signal, V_{AB}, inevitably occurs with the shape of a time derivative of $P(t)$ and, consequently, at output stage the integration of the signal must be performed in order to recover the original $P(t)$ waveform. Figure 3.12 shows the three topologies that can be used in the amplifying stage, immediately after the sensor. The VMA and CMA topologies have been studied by others, [170, 172], but little or no attention, has been regarded to the Active Differentiator Mode amplification (ADM) type. Their performances were evaluated by determining the Root Mean Square Error (RMSE) between V_{Test} and the final circuit output voltage (V_{out}), represented in Figure 3.10.

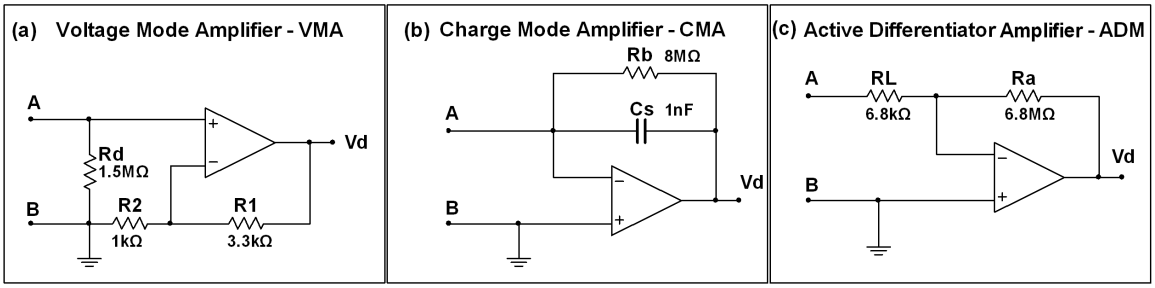

Figure 3.12: The three main topologies for amplifying the sensor signal: (a) voltage or passive differentiation mode amplifier (VMA), (b) charge mode amplifier (CMA) and (c) active differentiator mode amplifier (ADM).

3.3.1.3 Peak Detector and Timer

Cardiac waveforms show a pulsatile profile, characterized by a prominent peak and the purpose of this sub-circuit is to extract the reference time signal associated to this peak. To achieve this goal, V_d is fed to a peak detector formed by two peak stretchers ($U2$, $D1$, $C2$, $R4$ and $U4$, $D3$, $R5$, $C3$) a level shifter ($U3$, V_{ref}) and a comparator ($U5$) (see Fiqure 3.10). To avoid false triggers the peak stretchers have different time constants, $R4C2$ and $R5C3$, chosen according to the values of typical signals, and one of them is level shifted before being fed to the comparator. The signal at $R4$ is slightly negatively shifted by the instrumentation amplifier $U3$ at a gain of 1 ($R3 = \infty$) using an adjustable reference voltage, V_{ref}, derived from the supply voltages (not shown). The negative shifting, although it introduces a small delay, is essential to avoid erroneous triggering of comparator $U5$ (LM311). In fact, without the shifting action of $U3$, the two signals at the comparator input would remain close enough to cause repeated triggering of the comparator during the rising edge of the signal, even for low noise levels. Physiologically, this instant occurs when the rate of change of the systolic pressure upstroke is maximum, a few milliseconds prior to its peak and, it is unambiguously determined for a large number of differently shaped real and synthesized waveforms. In practice, the output signal exhibits a SNR higher than 40 dB and peak amplitudes of around 1.5 V, allowing a very clear peak detection operation exhibiting very low levels of jitter.

The comparator output, in the DAQP1, triggers a tandem of two timers. The first one is set to produce an up to 1.2 s wide pulse that acts as a delayed trigger to the second 1.5 ms wide reset pulse used by the integrator block. The timing sequence, based on two LM555 timers is represented (details omitted) by the box labelled Timerx2 in Figure 3.10. In Figure 3.13 this process is detailed

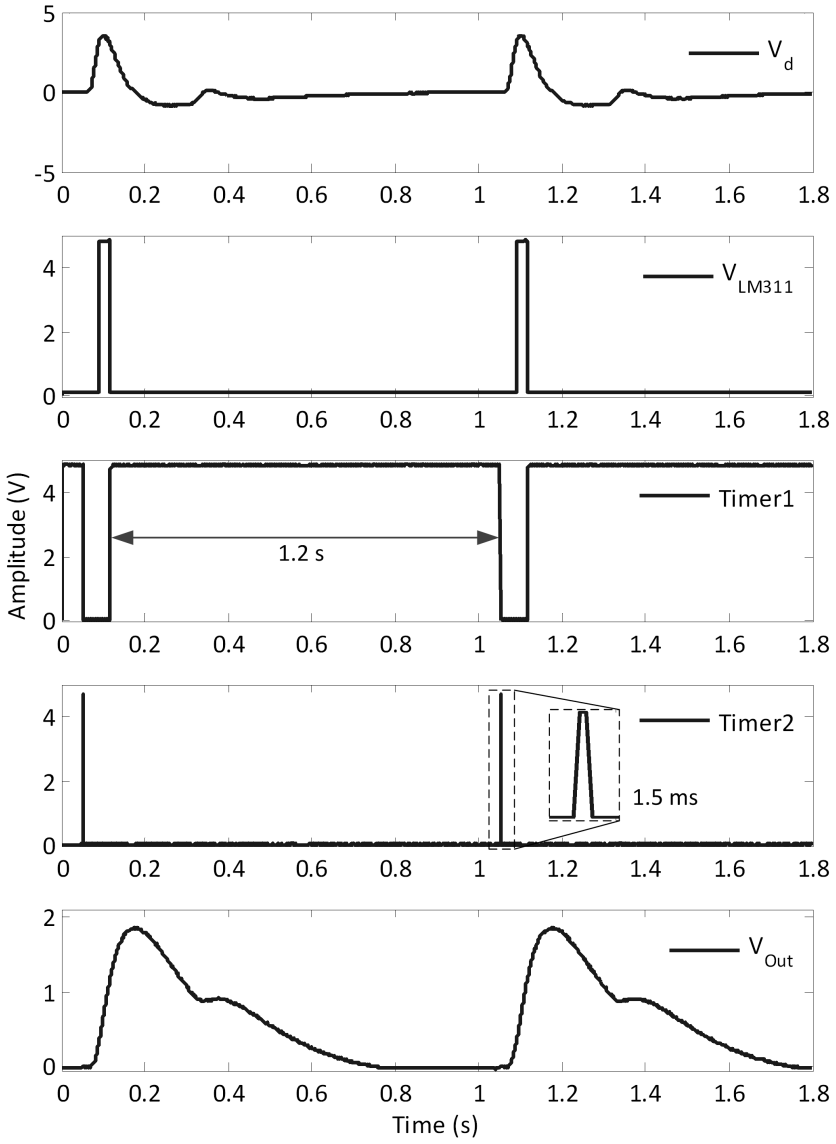

Figure 3.13: Amplitude and phase of the transferisk function of the circuits. Vertical axes are in decibels (dB) for amplitude and degrees for phase.

using simulated data: the V_d signal that is fed to the peak detector circuit, the identified peak at the LM311 output, and the Timer1 signal that precedes the Timer2 signal, also called, the RST signal, that actuates before each one of the pulses.

3.3.1.4 Integrator with Reset

A precision integrator ($U7$, $C4$ and $R6$) was implemented in the DAQP1 to ensure the re-establishing of the original input signal. The integrating current was set by $R6$ (10 $k\Omega$) to a maximum value below

0.1 *mA* for a typical 1 *V* peak in V_d, a value well within the current capability of the OPA2704 output.

It should be noticed that the timing of the reset signal, delivered to the DG419 analogue switch, described in the previous sub-section, leaves the physiologically interesting parts of the cardiac signal (mostly the systolic upstroke, systolic peak and diastole) untouched and is compatible with a wide range of HR values.

Charge injection, 10 *pC* maximum for the DG419, shows up only as a few microVolt voltage step due to the large value of *C*4 and, more importantly, it affects the cardiac pulse waveform at its very end, a region of little physiological interest. This electronic block was replaced in DAQP2 and DAQFP (Section 3.3.2.2) by the inclusion of a microcontroller.

3.3.1.5 Baseline Restorer

This concept of BLR, introduced by Robinson [173] is used in nuclear detector electronics, where the baseline of detector pulses generally is randomly shifted due to the combined effect of the unpredictable time of arrival of particles and the capacitive coupling of the detector signal, necessary to block its high voltage bias. All AC-coupled amplifier systems in pulsed signal applications suffer from baseline variations, which originate errors in measurements that depend on the accuracy of the pulse amplitude measurement. The capacitive nature inherent to the PZ sensor front-end, inevitably turns the system into an AC coupled circuit, subject to strong baseline deviations that, in addition to other errors, impairs real time visualization of the signals. The elimination of the baseline drift by the BLR circuit consists in forcing the foot of systolic pulses to start close to zero without affecting the shape of the signal. This requirement is met if the BLR is activated only at the end of the cardiac pulse, just before the upstroke of a new pulse.

Like many other biosignals, APW pulse drift essentially correlates to three sources: respiratory activity, variations in signal shape and signal jitter (defined as a random variation of the pulse period). If an operator holds the probe during data collection, an extra source of baseline drifting shows up due to the variations of its interaction with the tube of the phantoms or, in the case of biomedical applications, with the patient. The use of a fixed support, a collar in biomedical applications, rules out the influence of the operator. In biomedical applications, the low frequency respiration effect,

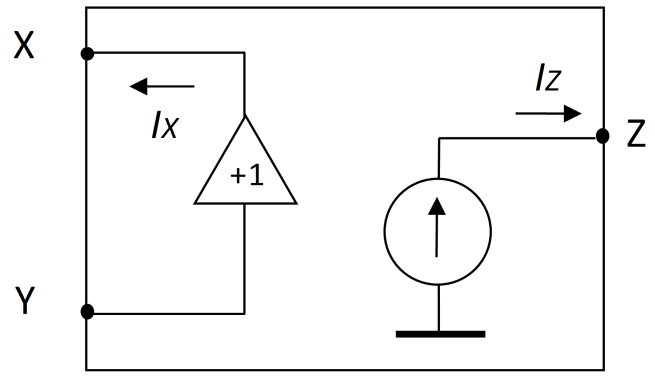

Figure 3.14: Equivalent electronic circuit for the *CCII*+.

although attenuated by the AC coupling itself, and even if it is minimized by the use of a collar, shows up, at the output, as a major contributor to baseline drift. However, the last two sources (signal shape and signal jitter) are intrinsic and can only be eliminated by the BLR circuit.

A BLR circuit was developed based on the properties of the Second Generation Current Conveyor (CCII) OPA860, referred, by the manufacturer (Texas Instruments®), as an Operational Transconductance Amplifier (OTA) [174]. The defining equations are described in Equations 3.1 and 3.2 and, the CCII symbols are depicted in Figure 3.14. The signal following the acronym CCII (CCII+ or CCII−) refers to the relationship between the direction of *IX* and *IZ*. In the OPA860 case, CCII+ means that the two currents are mirrored. Unlike operational amplifiers, the CCII does not require the establishment of negative feedback to operate.

$$V_x = V_y \quad (V_x \; follows \; V_y) \tag{3.1}$$

$$I_z = I_x \quad (I_z \; mirrors \; I_x) \tag{3.2}$$

The OPA860 still offers an extra unity gain voltage buffer that can be useful in applications where its high input bias current (up to 8.5 μA) can be tolerated. It is not the case in this circuit, which mandatory demands the use of U6 (Figure 3.10).

Figure 3.15 explains the working principle of the CCII+ based BLR: when voltage at A tends to be negative, at the falling edge of each pulse, current $I_R = V_A/R$ flows through R and is mirrored

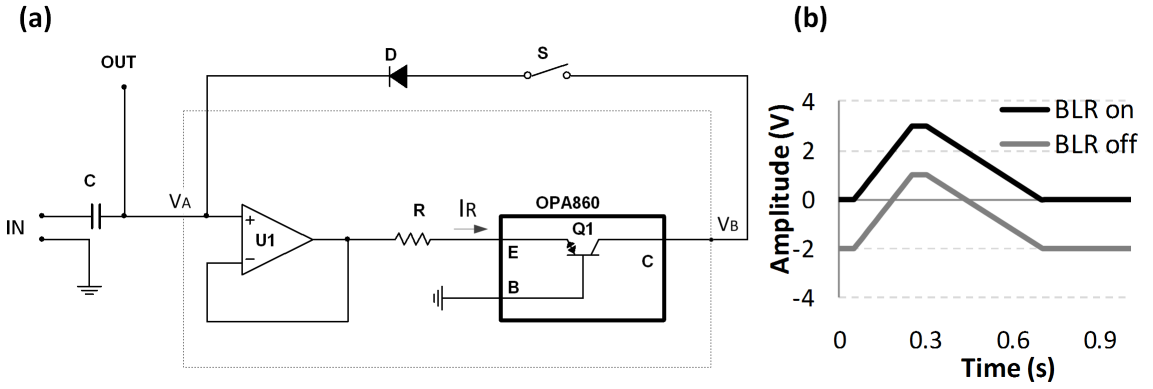

Figure 3.15: BLR circuit (a) and its Orcad simulation for a trapezoidal signal (b), with (S ON, black line) and without (S OFF, grey line) restoration.

at the B output. This linear unidirectional feedback mechanism, due to the action of diode D, tends to restore the baseline of the pulses in such a way that the foot of the systolic pulse always starts at zero, as depicted in timing diagrams of Figure 3.15(b), obtained by simulation. In the circuit of Figure 3.10, an adaptation of this principle is used. In this case, the BLR effect is obtained by the controlled injection of current from the CCII+ output in the summing point of the integrator, forcing the integrating capacitor to discharge, before a new pulse occurs.

3.3.1.6 Breathing Circuit

Respiratory rate is important for many clinical uses, including: detecting sleep apnea, sudden infant death syndrome, and chronic obstructive pulmonary diseases. The respiratory rate is also measured by many intensive care equipment [175]. To monitor respiration several methods can be used: thermistors, thoracic abdominal bands, nasal prongs and pneumotachographs. Thermistors are the most common sensors used for this purpose, in spite of showing a long time-constant response.

During several trials, it was verified that the breathing rate modulates both, amplitude and frequency of the APW signals, and that this information would be useful for further data processing. So, a circuit for breath rate monitoring was implemented, such as depicted in the APW module in Figure 3.14 (*Resp* dashed box). The circuit uses a Wheatstone bridge (see Figure 3.16), where R4 is a Negative Temperature Coefficient (NTC) thermistor.

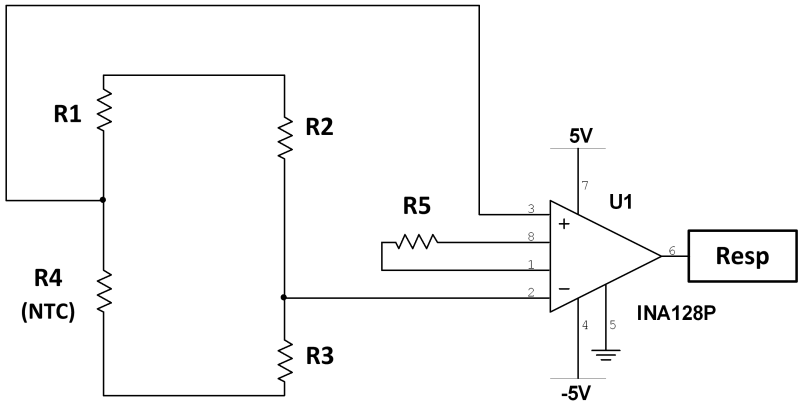

Figure 3.16: Breathing circuit implementation, where *Resp* is the respiratory signal.

3.3.1.7 Microcontroller Circuit

The advantages in developing a microcontroller based system show up in increasing the functionality of the system, as well as in decreasing the associated cost. There are different types of controllers technologies in the market. However, Digital Signal Processors (DSPs) assumed, over the last years, an increasing importance due to the enhanced functionalities that they are capable to grant. The development of this technology has enabled significant improvements in the speed analysis, in the accuracy, in the noise immunity, in the programmability, in the circuiting size reduction and, in addition, in the decrease of equipment cost. In particularly, the *dsPIC33* family from Microchip® employs a powerful 16-bit architecture that integrates the control features of a microcontroller with the computational capabilities of a DSP.

The adopted solution was dsPIC33FJ256GP710, that is a 16-bit, general purpose controller in a 100-pin DIL package, with a 40 MIPS processor speed, 256 $kbyte$ programme memory and a 30 $kbyte$ of RAM data memory. Figure 3.17 illustrates the complete circuit diagram. The blocks diagram, where the grey boxes outline the main hardware blocks, elsewhere described that communicate with *dsPIC* through the indicated ports. The blue blocks indicate hardware accessories required for the *dsPIC* operation. The microcontroller is clocked with a 8 MHz crystal and communicates with the host PC though the USB at a baud rate of 460800 $bits/s$. The microcontroller was programmed using a C-compiler from Microchip® (MPLAB ICD2), that is described in Section 3.3.2.2. Additionally, a push-button was added to allow microprocessor reset, as well as, a set of LEDs for testing and debugging purposes.

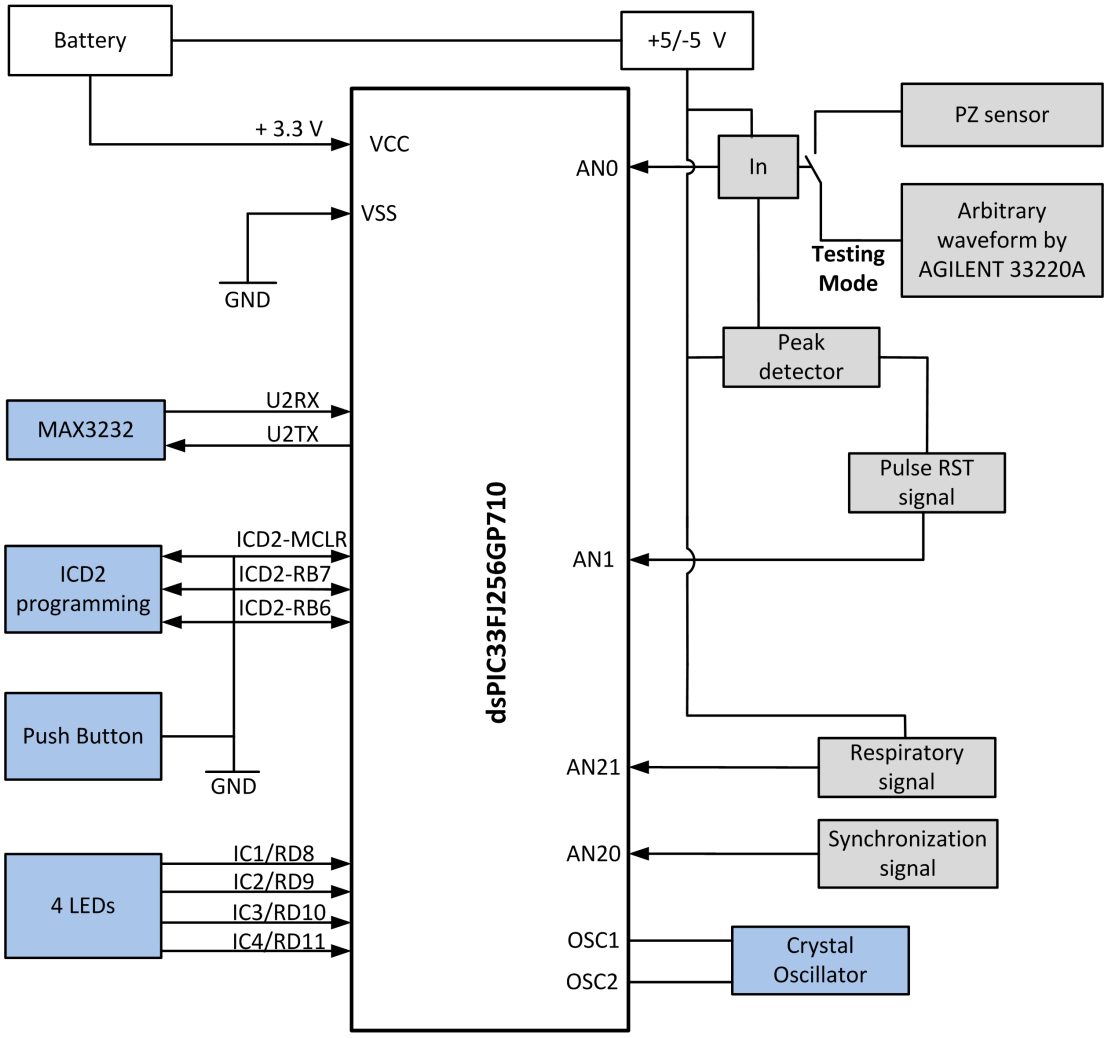

Figure 3.17: Generic block diagram of the hardware connections in the APW DAQ board. The grey boxes indicate the main hardware signal blocks that communicate with *dsPIC*, and the blue blocks the hardware accessories required for the *dsPIC* operation.

3.3.2 Firmware Development

Several cardiovascular applications have been previously reported in the literature, using microcontrollers: Klig *et al.* [176] used these systems for BP monitoring and ECG signals; Bing-Nan *et al.* [177] proposed an embedded medical advisory system for mobile cardiovascular monitoring devices, that provides microcirculation information; while Germano *et al.* [178] introduced a generic architecture for biomedical embedded systems, dedicated to clinical analysis and patient monitoring.

There are two important parts of design procedure: the modelling and simulation of the system

dynamics and the implementation of the control algorithm within the microcontroller. The modelling task was implemented in the Matlab/Simulink platform. This option allowed to speed up the prototype development cycle, since programming the microchip-embedded digital controller directly from a Simulink model, significantly speeds up the process.

3.3.2.1 Pulse-by-pulse Control

The pulse-by-pulse control strategy, which is presented here, grants continuous and stable real-time monitoring. It is important to ensure the capability to deal with a heterogeneous set of patients, despite any disturbances suffered by the body, such as patient movements or respiration rate. The main control scheme includes the baseline detection and pulse-by-pulse reset based on the peak-detector signal described in the Section 3.3.1.3. The implemented control algorithm allowed the advantageous replacement of the implemented hardware in DAQP1, which comprised a set of two timers that were not able to deal with the physiological variability between subjects.

As previously mentioned, the signal from the peak-detector circuit corresponds to the highest peak rise slope in the PZ signal, that occurs a few milliseconds after the systolic upstroke, denominated t_{min}. To deal with this constraint a delay-time needs to be introduced to force the correspondence of the beginning of the integration process, marked by the peak-detected signal, with the beginning of each pulse. The determination of the exact time delay required is not possible, so, alternatively, the estimation of this value (t_d) is used from previous physiological statements. Firstly, it is known that the typical values associated to the upstroke time (t_u) vary between 110 *ms* and 230 *ms*, as determined by Buteler [179]. It is also known that t_u values are typically twice of the t_{min} value that is required for the algorithm. Therefore, t_{min} is, approximately, in the range of $55-115$ *ms* (Equations 3.3 and 3.4). During the delay computation, it is preferred to set this value near to the upper limit to avoid errors in low upstroke time values, that are commonly associated to more complicated physiological situations. So, a t_d value of 110 *ms* was adopted.

$$t_u = 2 \times t_{min} \tag{3.3}$$

$$55 < t_u < 115 \tag{3.4}$$

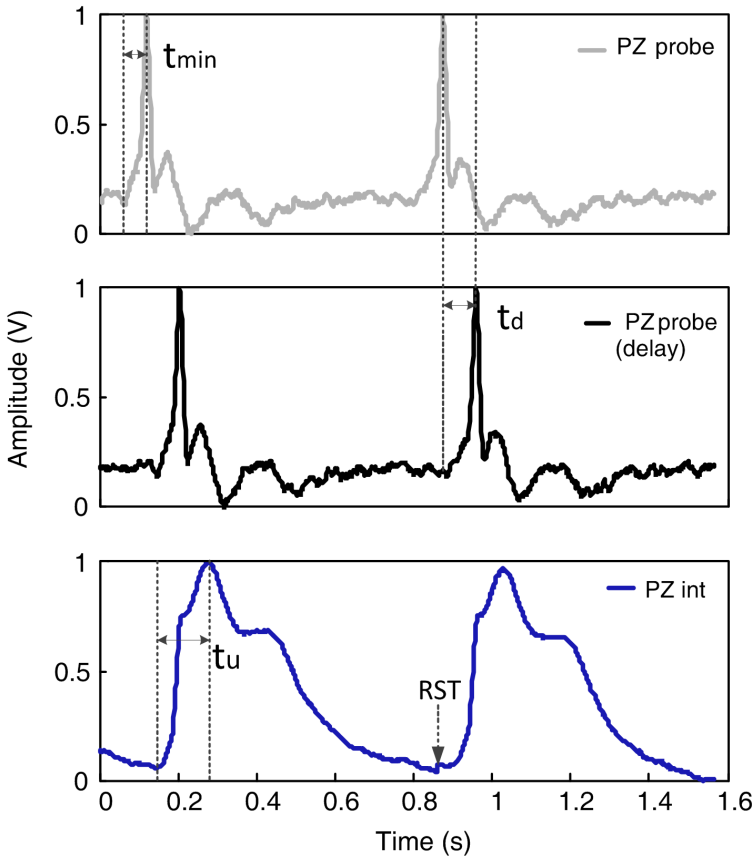

Figure 3.18: PZ waveform is represented in grey line and delayed PZ waveform in black line. The integrated waveform PZ int (integrated) represented in blue line was obtained from the delayed PZ waveform. The t_{min} represents the ideal delay value, t_u is the systolic upstroke time and t_d is the delay time used in the PZ waveform integration.

Figure 3.18 shows the original PZ signal (grey line) and the delayed PZ signal (black line). Integration (blue line) is performed for the delayed PZ signal. As the pulse-by-pulse analysis requires a reset pulse, a special concern with baseline is also required in this process. Reset signal (RST) sets the signal to 0, and as the signal for the *dsPIC* is monopolar, the mean value is of the order of 1.5 V. This situation leads to the occurrence of drastic level decrease at begin of each pulse. To overcome this limitation, the mean value is removed internally.

This mechanism is essential for the real time pulse-by-pulse analysis. However, an alternative post-processing algorithm (see Section 4.4.2) was also developed to further processing raw data, avoiding some inaccuracies that can occur during real time pulse segmentation.

3.3.2.2 Simulink Implementation

The Simulink model presented in Figure 3.19 intends to be an alternative to the hardware described in Section 3.3.1.5. Its implementation is based on the available blockset *Embedded Target for PIC/dsPIC* [180]. The *ADC Input block* converts the analogue signals from the PIC input channel, configured to work as 10 bit ADC converter, according to the specifications. Then, in the Simulink, this value is converted in a signed 16 bit variable (*Data Type Conversion*), where just the 10 lower bits are used. The *Tx Output Multiplexed for Matlab-Labview Output block* can write 16 variables to Matlab® using the serial protocol. In case the host computer has no serial interface, an USB to serial converter hardware is used. Configuration options are detailed in Appendix B.

In a typical measuring session, the first few pulses are used just for system adjustment, mainly HR measurements. Then, real data collection lasts for as long as possible, without causing discomfort to the patient, who is asked not to swallow during data acquisition. Typically, less than two minutes allow for the acquisition of a number of cardiac pulses with enough statistical relevance.

The state-of-the-art acquisition systems, with similar purposes, hardly exceed acquisition rates of 300 Hz. In this work, a bandwidth of 500 Hz was considered as adequate, since the relevant

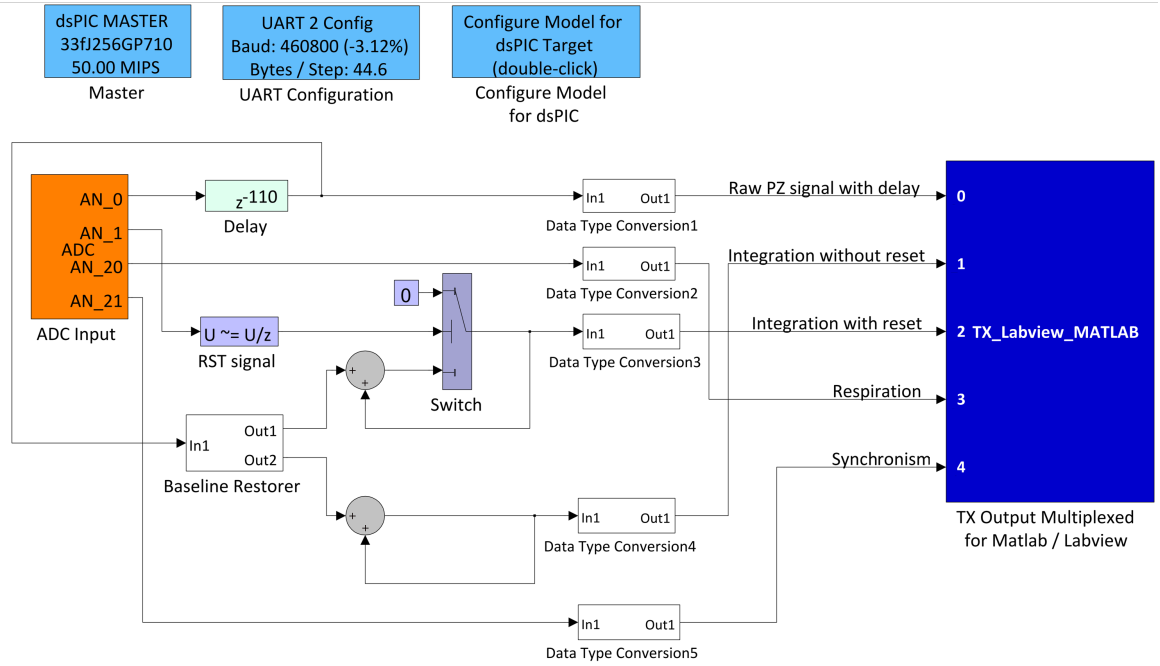

Figure 3.19: Simulink model used for C-code generation for *dsPIC* using Kerhuel blockset [180].

information remains below this value, such as presented in Section 3.4.1, and consequently a 1 *kSPS* acquisition rate was used throughout

3.3.2.3 *dsPIC* Programming

Microchip® MPLAB v8.30 was used in the models programming, through the C30 compiler, and using the C-code generated by the Simulink model. This approach provides an interactive graphical environment where the algorithms are developed in the form of block diagrams. Simulink platform also allows the interface required to real time data transmission through Universal Asynchronous Receiver/Transmitter (UART). In Figure 3.20 is shown a flowchart with the main stages of programming.

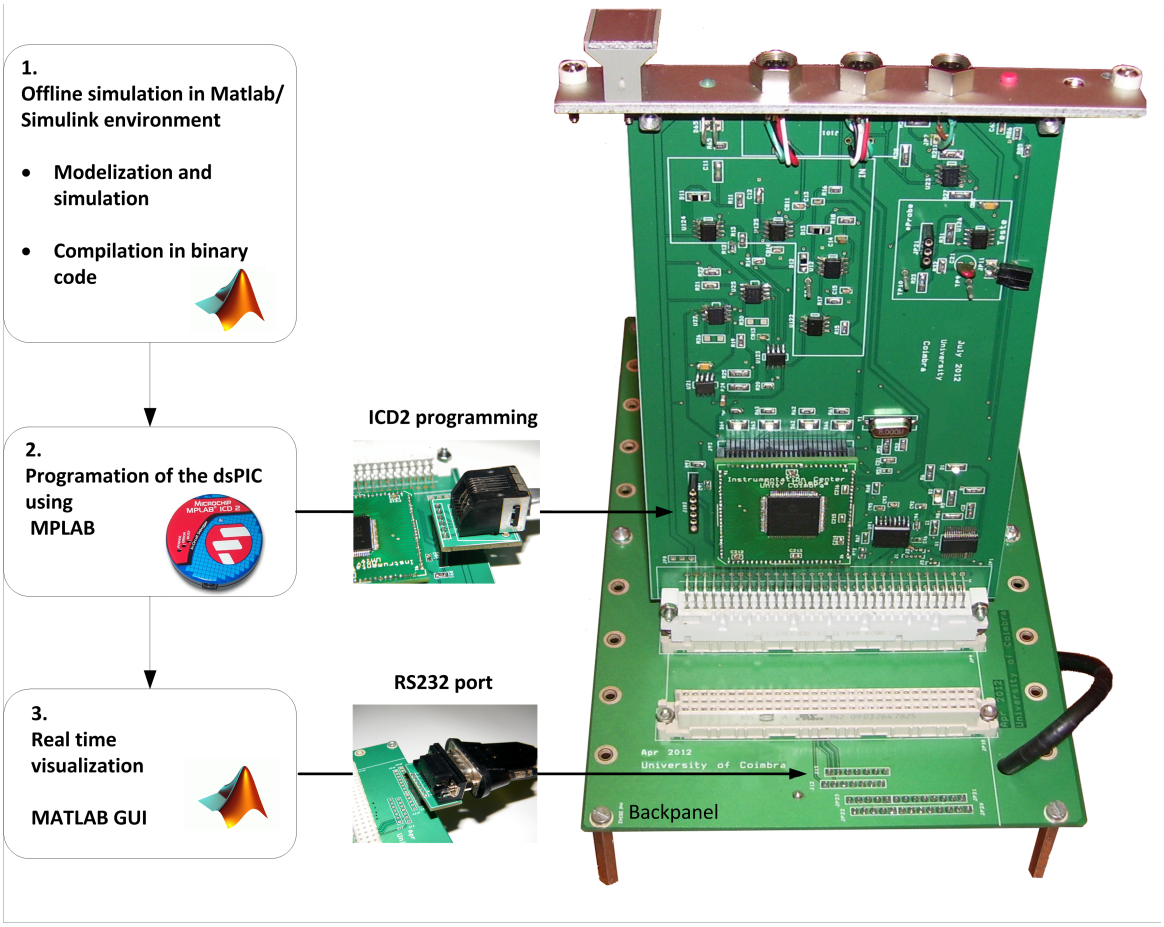

Figure 3.20: Schematic of the stages of programming. (1) The model is generated in the Simulink platform and, then, is compiled in C code; (2) programming of the *dsPIC* using the MPLAB ICD2; (3) microcontroller real time data transmission through UART.

3.3.3 Probe Mechanical Design

The distension wave associated to the propagating of the pressure wave, generates forces that, given the adequate mechanical interface transmission mechanism, actuate upon the sensor. This mechanical interface plays an important role in the overall performance of the probe, as it determines the forces that are exerted over the sensitive part of the sensor body. Apart from the PZ sensor itself (brass coated disc), the main mechanical skin contact element of the probe is the mushroom-shaped part shown in Figure 3.21(a) and (b), responsible for transmitting the distension associated to the pressure wave. The transversal and shear effects are suppressed and only the longitudinally external applied forces are allowed to originate an electrical response. The probe dimensions are depicted in Figure 3.21(c) and the final probe is depicted in Figure 3.21(d).

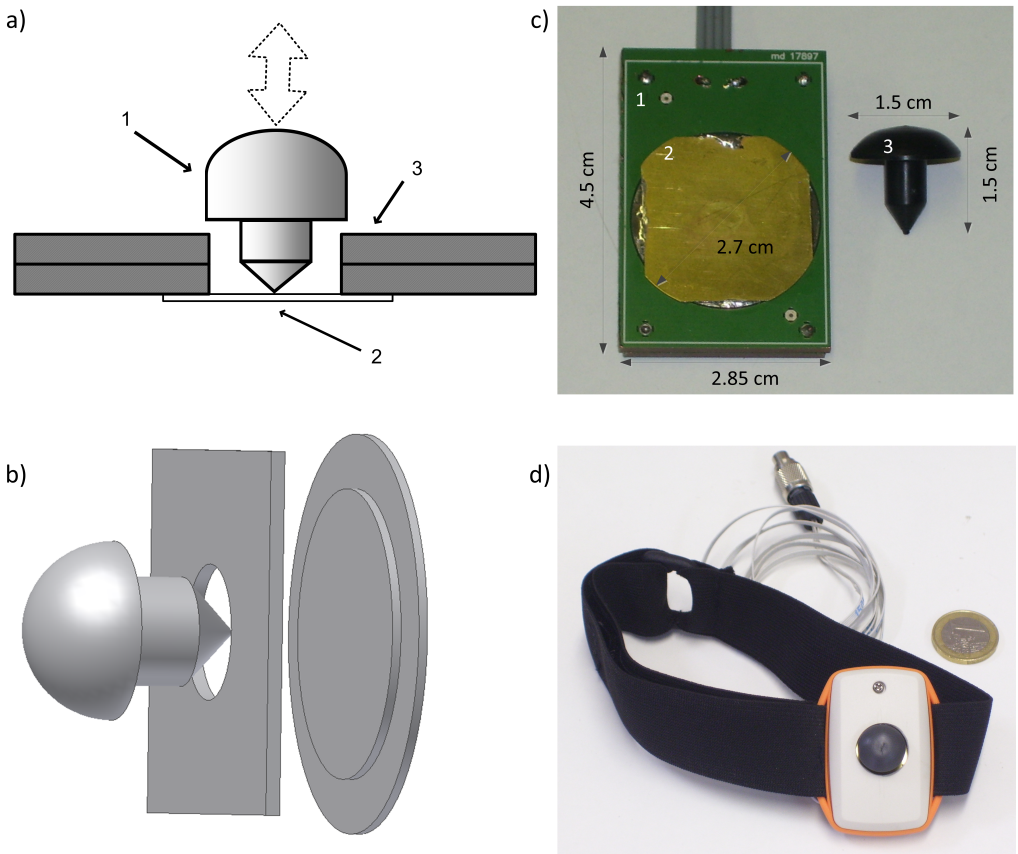

Figure 3.21: Mounting details of the probe elements shown in a transversal cut (a) and 3D exploded (b), views. The arrow in (a) indicates the externally applied forces and, (1) is the mushroom-shaped interface, (2) is the PZ disc sensor and (3) is the PCB frame. In (c) it is shown the upper view and probe dimensions, where (1) is the support PCB layer, (2) is the PZ disc and (3) the mushroom-shaped PVC (PolyVinyl Chloride) interface, and in (d) is shown a photo of the probe with collar.

3.4 Performance Tests

The evaluation of the performance of the APW electronic board was accomplished taking into account functional tests.

3.4.1 Spectral Bandwidth

The spectral limit of the physiological information contained in the APW pulse is an important issue. This information is essential for the processing routines that can include filtering procedures. The frequency components of cardiac waveforms were analysed by using the FFT routine, as it is shown in Figure 3.22 and, it is possible to conclude that the critical information spans up to a $20-25\ Hz$ higher limit.

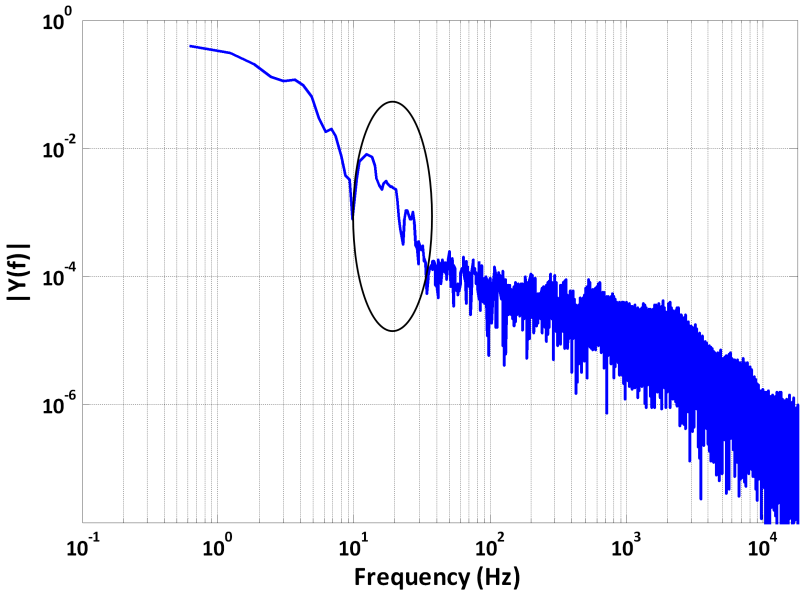

Figure 3.22: Harmonic contents of the amplitude response of an APW segment, where is pointed the interest zone. Frequencies higher than 25 Hz are within in the noise plateau.

3.4.2 Analogue Circuit Transfer Function

Figure 3.23 shows the amplitude and phase of the transfer functions of the full circuit, for each of the amplifying circuits of Figure 3.12, obtained using *Pspice* simulation. The grey and black lines

refer to the ADM circuit and, clearly, lead us to expect a very good performance in terms of RMSE due to its uniformity along frequency, for gain and phase.

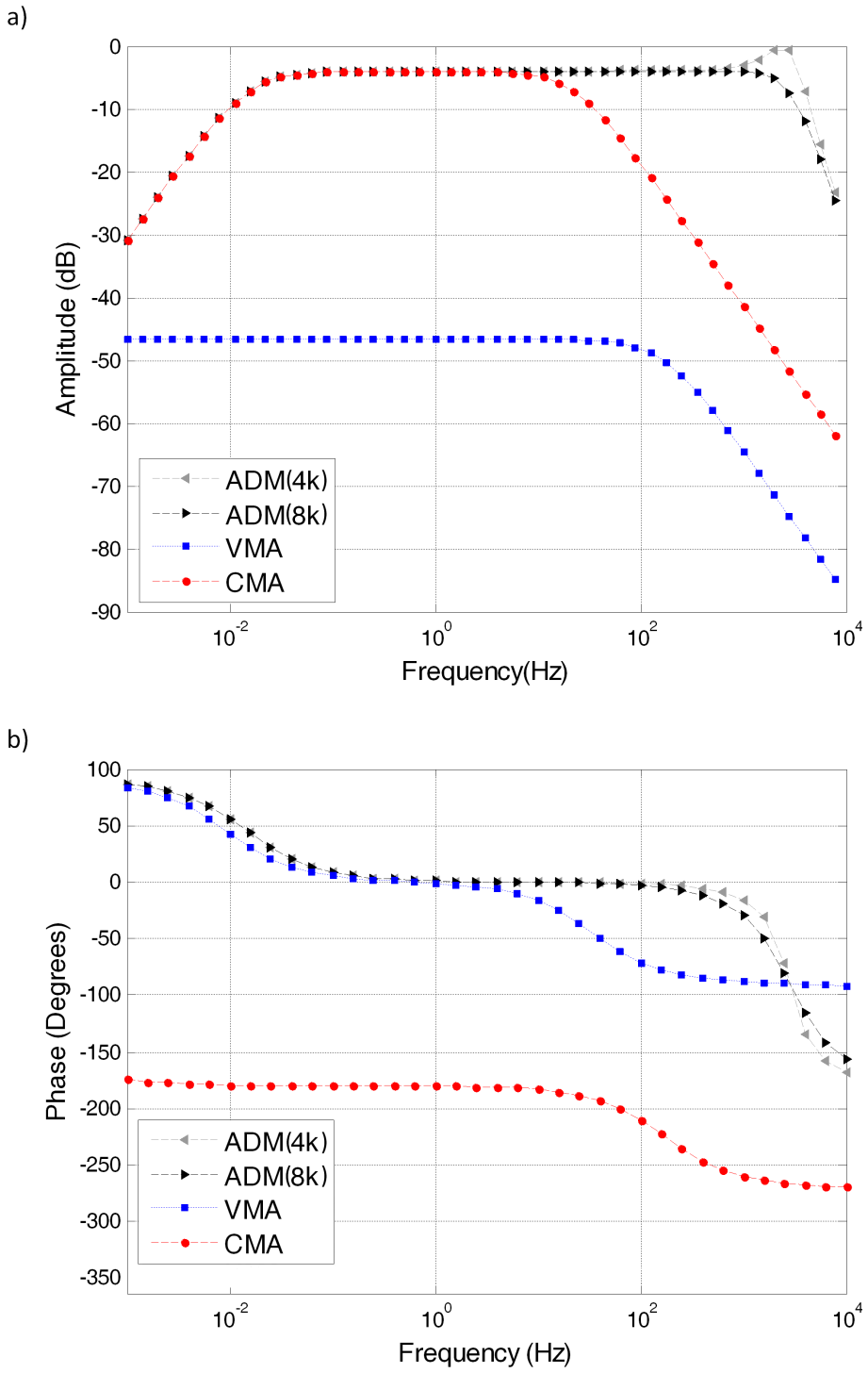

Figure 3.23: (a) Amplitude and (b) phase variation of the transfer functions of the ADM(4K), ADM(8K), VMA and CMA circuits.

At the high frequency end of the ADM curve (without resistor R_L), the effect of Common-Mode Rejection Ratio (CMRR) degradation becomes visible and shows up as a sudden as a sudden increase in gain (grey line). The CMRR *versus* frequency response curve generally holds constant for low frequency values and starts to decay at a certain frequency value. If unattended, this CMRR spike would allow the amplification of high frequency noise which would greatly impair the signal. The resistor R_L (see Figure 3.12(c)) counterbalances this effect, by cutting the high frequency response of the circuit. The frequency band response of this circuit is limited, according Equation 3.5, yielding the behaviour shown by the black line. This effect imparts the important feature of preventing the high frequency noise to contaminate the signal, as it is very common in differentiating circuits, and, consequently, impairing the extracting processing feature, that depends on it. The value of R_L is determined by the value of the selected cut frequency combined with the equivalent sensor capacity C_S, 2 kHz and 8 nF, respectively, according to:

$$f = \frac{1}{2\pi R_L C_S} \tag{3.5}$$

In Figure 3.23(a), in the amplitude response, the effect of R_L is stressed by showing the two responses obtained from a well balanced (black line) and an under balanced circuit (grey line) corresponding to R_L values of 8 $k\Omega$ and 4 $k\Omega$, respectively. The VMA curves reveal very low gain, as well as a great deal of phase shift (Figure 3.23(a) and (b)) respectively.

As the frequency range of cardiac signals spans up to around 25 Hz (Figure 3.22), it is not surprising that both, ADM and CMA, show very good amplitude response and phase uniformity, because their operation is based on an operational amplifier circuit that behaves very close to ideal, for this frequency range. In principle, the CMA circuit could be used with no further integration to delivering the final cardiac waveform. However the availability of a differentiated signal is essential for extracting the time reference signals used by the BLR and by the integrator. Therefore, a configuration where the ADM circuit is followed by the integrator was adopted, which, in any case, is the best performing configuration, as it is presented in the next section.

3.4.3 Input/Output Error

A set of cardiac-like synthesized pulses [181] shaped to reproduce the different classes typified by Murgo [182], were used during the RMSE calculation for simulated and real data. In the real data tests, the test input (V_{Test}, represented in Figure 3.10) is driven by the arbitrary waveform generator (excited with the same signal type used in the simulations). RMSE is calculated taking V_{Test} as the reference, thus

$$RMSE = RMS(V_{out} - V_{Test}) \tag{3.6}$$

If any of the circuits performs ideally (no distortion introduced) its RMSE would be zero, in the absence of noise. However, if noise is present, the ideal circuit yields a non-zero, minimum value, for RMSE:

$$RMSE_{min} = \sqrt{n_{VS}^2 + n_{Vout}^2} \tag{3.7}$$

where n_{VS} and n_{Vout} are the RMS of noises at the input and output, respectively. The $RMSE_{min}$ has been calculated to all waveforms, that were previously normalized to discard interference of gain and offset, that are not important for evaluating the accuracy in reproducing the input waveform. Finally, the so called $RMSE_c$ is computed, subtracting the noise contribution from RMSE according to,

$$RMSE_c = RMSE - RMSE_{min} \tag{3.8}$$

In Figure 3.24 a Type B (see Section 4.4) waveform collected at V_{in} and at V_{out} is shown for each one of the electronic topologies at study. The simulated and measured data are represented in (a) and (b), respectively. The RMSE values between input and output for each of the three topologies, computed from simulated as well as measured data, using the very same synthesized pulse waveforms as inputs in the simulator and in the physical circuits, are shown in Table 3.3, for all four types of cardiac waveforms. The circuit data is split in three columns to show the contribution of noise. The results clearly demonstrate the superiority of the ADM topology in the

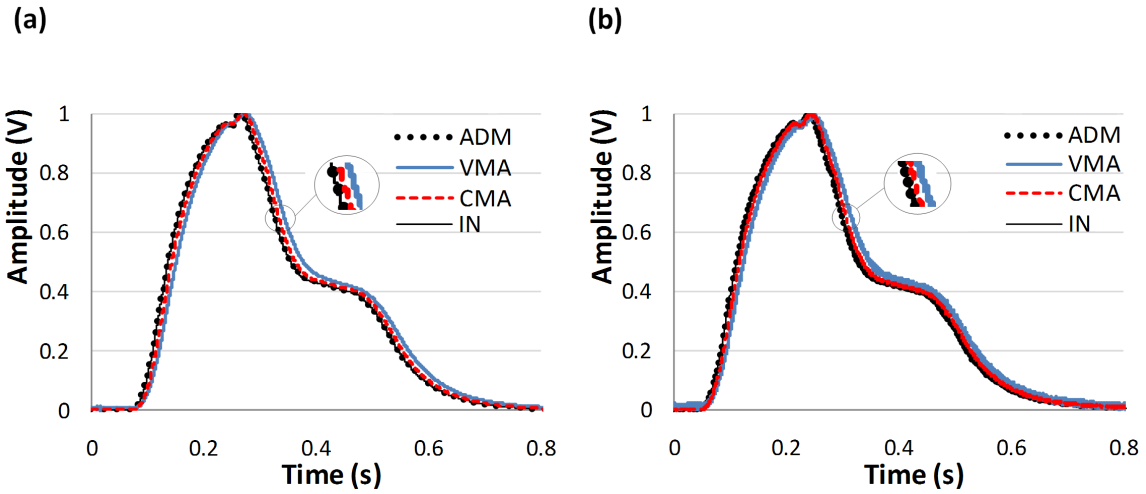

Figure 3.24: Normalized input and output for a Type B cardiac pulse in the three circuit topologies, as given by the simulator (a) and measured in the circuit (b). Noise is only visible in the real data, plotted in (b).

precise rendering of the input signal waveform. Therefore, the ADM topology was used to perform the mechanical tests, as well for the test bench characterization and the *in vivo* tests.

Table 3.3: RMSE measurements for each of the three amplifying circuits, using cardiac - like synthesized pulses to reproduce the different classes described in the literature. Data are taken from simulator (PSpice) and measured in the circuit (real data).

Circuit	**Wave Type**	**Simulation**	**Real Data**		
		RMSE (%)	*RMSE (%)*	$RMSE_{min}$ *(%)*	$RMSE_c$ *(%)*
VMA	A	5.53	4.39	0.05	4.34
	B	5.36	4.65	0.04	4.61
	C	5.68	3.66	0.05	3.61
	D	5.32	3.47	0.05	3.42
CMA	A	2.52	2.27	0.05	2.22
	B	2.47	2.11	0.06	2.05
	C	2.61	2.13	0.05	2.08
	D	2.34	2.04	0.06	1.98
ADM	A	0.14	0.31	0.31	≈ 0
	B	0.19	0.33	0.33	≈ 0
	C	0.07	0.33	0.33	≈ 0
	D	0.14	0.30	0.30	≈ 0

3.4.4 Pulse Delay

Due to the specificity of the described module (the one that makes use of the PIC in the multichannel platform), signals coming from the PZ transducer are delayed, and as consequence some computations are made. On the contrary, the synchronization signal described in Section 3.2.4 is acquired and sent to the PC with no firmware transformations. Therefore, it is extremely useful to measure the delay between signals from the APW module and synchronization module to assure that signals are correctly time aligned.

The time delay was measured using the same signal fed to two different ADC channels of the PIC. One used as the synchronization channel, fed directly to the ADC, and another following the normal PZ path, since the V_{Test} input (Figure 3.10) to a ADC channel. For each channel, a pulse train with a frequency of 1 Hz and a duty cycle of 20 % was used as the input signal. Signals were generated using two Agilent 33220A, acting, one as a master, and the other one as a slave, so that external trigger could be used in the slave to synchronize both signals. Both were synchronized with a mismatch of 1 μs, which was assessed with the oscilloscope. This error is negligible, given the time scale that is used in signal processing routines.

Figure 3.25 shows both signals, being visible the delay between both signals. The delay was computed from differentiated signals, where the beginning of each pulse (and also its ending) are clearly identified by peaks, as it is visible in Figure 3.25(b).

A total of 19 acquisitions, comprising 1712 variable length valid pulses were performed. Between each acquisition, power was turned off, so that the PIC could be reset. A delay value (t_{delay}) of 112 *ms* was measured, which is close to the expected value, since firmware routine introduces a delay of about 110 *ms* (see Section 3.3.2.1).

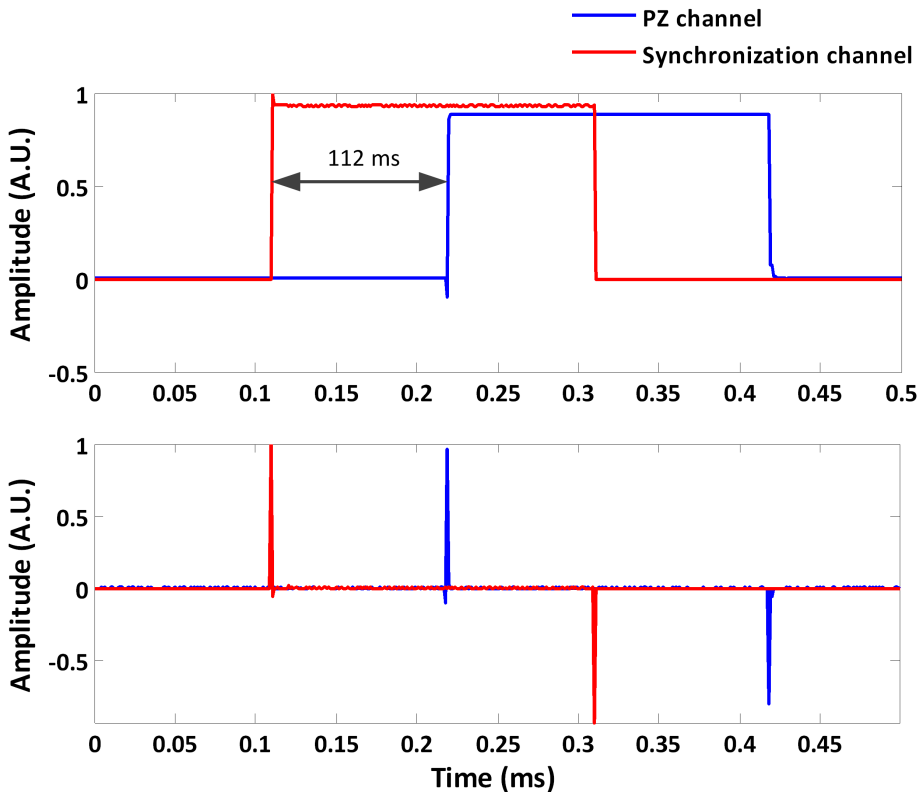

Figure 3.25: In (a) the pulses acquired from both ADC channels are shown, where the delay between channels is clear and in (b) the pulse differentiation of both ADC channels is depicted.

3.4.5 Optimal Mechanical Design

A number of tests were conducted to determine the optimal shape of the *mushroom* mechanical front-end interface in order to repeatedly obtain the best (undistorted) pressure waveform transmitted to the sensor. In these tests three, differently shaped, *mushrooms* were fitted to the probe and submitted to the direct action of the actuator. The difference is in the geometry of the *mushroom* base, which directly contacts the sensor, that can be shaped pointy, flat and intermediate. The stroke of the actuator was adjusted to about 10 μm, and the response was collected by the probe and stored in the memory of the Data Acquisition System (DAS). Again, the overall error is quantified as the RMSE between normalized input and output, V_{Test} and V_{out} (see Figure 3.10), respectively. Results, shown in Figure 3.26 and Table 3.4, are obtained for an Type A (see Section 4.4) cardiac waveform [181], because it is the most demanding in terms of circuit performance. The pointy probe exhibits the best performance in reproducing the input waveform with the lowest RMSE variance.

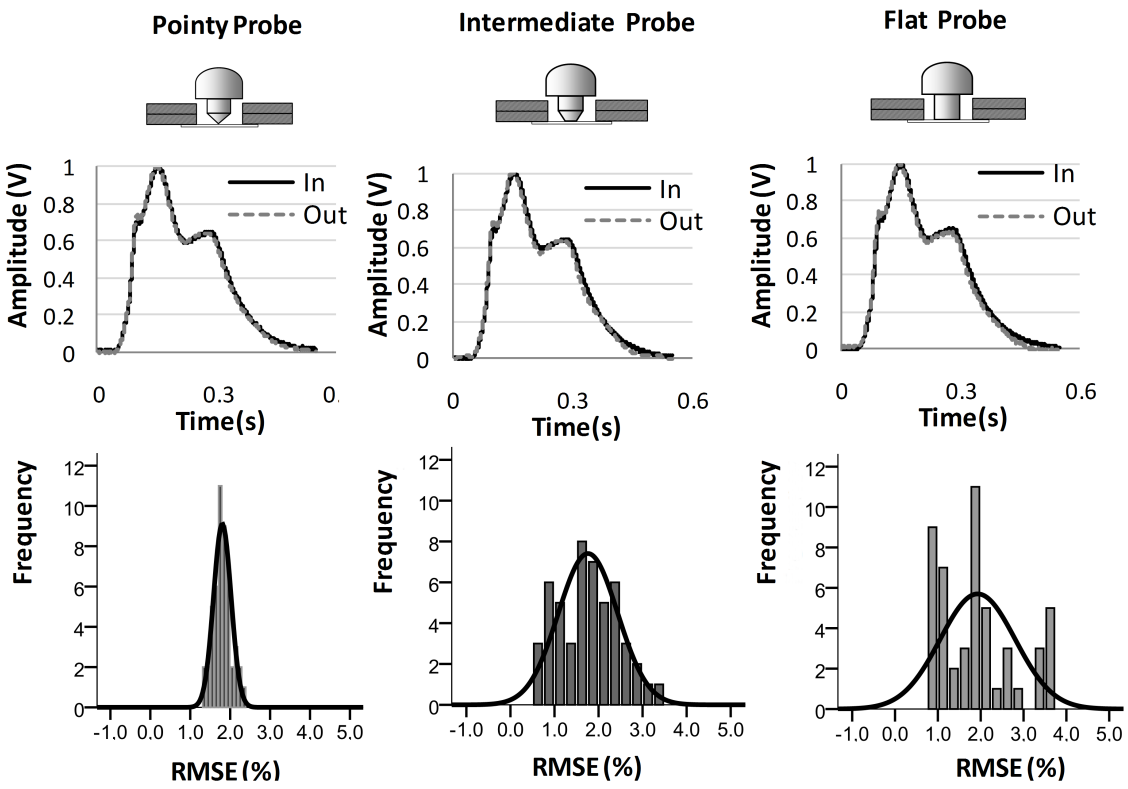

Figure 3.26: RMSE for differently shaped sensor interfaces measured for a Type A waveform. In the first row, the three probe geometries are diagrammatically shown. The second row depicts the input and output acquired signals. The last row shows the RMSE distribution, demonstrating the superior consistency of the pointy probe.

Table 3.4: Statistics of measurements depicted in Figure 3.26.

Probe type	**Pointy**	**Intermediate**	**Flat**
Number of pulses	50	50	50
Mean	1.80	1.76	1.93
SD	0.22	0.67	0.88
Minimum	1.34	0.66	0.77
Maximum	2.30	3.28	3.62

3.5 Test Bench Characterization

For testing the probe, as well as the several parameter extraction algorithms, a special purpose test bench was developed [96], represented in the diagram of Figure 3.27. The pressure wave was generated by a piston mechanism coupled to a 0.7 *mm* stroke actuator (ACT) driven by a high voltage

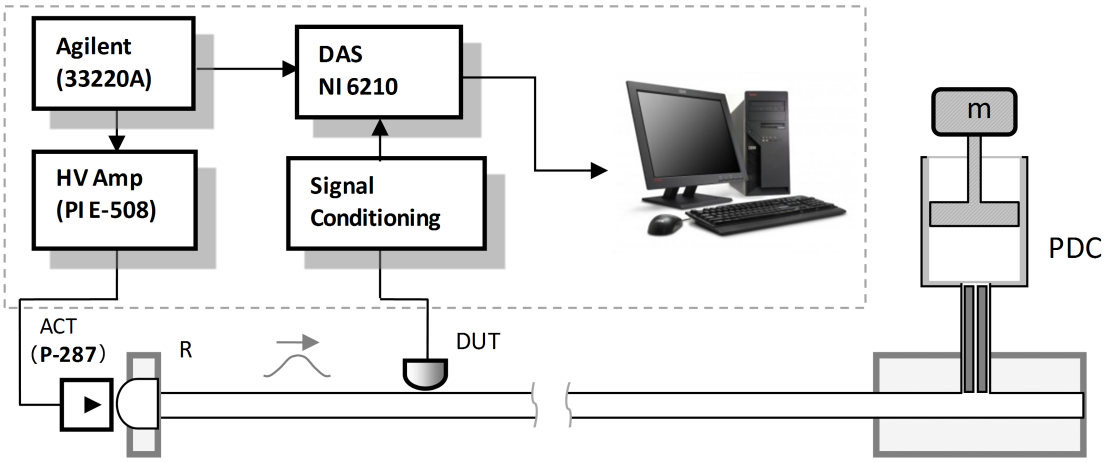

Figure 3.27: Schematic representation of the test bench. The signal conditioning block encloses the circuit of Figure 3.10 and the device under test (DUT) is the probe presented in Figure 3.21.

linear amplifier (HV Amp) (Physik Instrumente GmbH, ref. P-287 and ref. E-508, respectively) and launched into an 8 *mm* internal diameter, 0.5 *mm* wall thickness silicone rubber tube (from J. Lindemann GmbH). The waveforms were programmed into an arbitrary waveform generator (AWG) Agilent 33220A. According to the HV Amp specification, for the 290 *nF* load represented by the ACT, a 200 Hz bandwidth can be obtained if the output of the E-508 does not exceed 200 V. Such bandwidth allows the synthesis of the more sophisticated cardiac-like waveforms, as well as low frequency triangular, sinusoidal and Gaussian test signals used to characterize the test bench. To comply with these constraints, as the gain of the E-508 HV amplifier is factory fixed at around 100, the amplitude of the waveforms at the AWG output does not exceed 2 V. In addition to the propagating pressure wave, a static pressure level (PDC) of around 60 *mmHg* is forced by a piston-mass (m) combination at the other extremity of the tube.

This test bench allows the full study of the propagating wave dynamics, including reflections at the extremities of the elastic tube, as is visible in Figure 3.28. The forward and the repeatedly reflected waves, originated by a Gaussian shaped stroke of 100 *ms* duration along the tube, can be clearly seen. The map of Figure 3.28(a) is obtained for successive sensor positions, 2 *cm* apart, covering the full 80 *cm* of tube length. Figure 3.28(b) shows one single pressure curve collected by the probe positioned at 48 *cm* from the ACT.

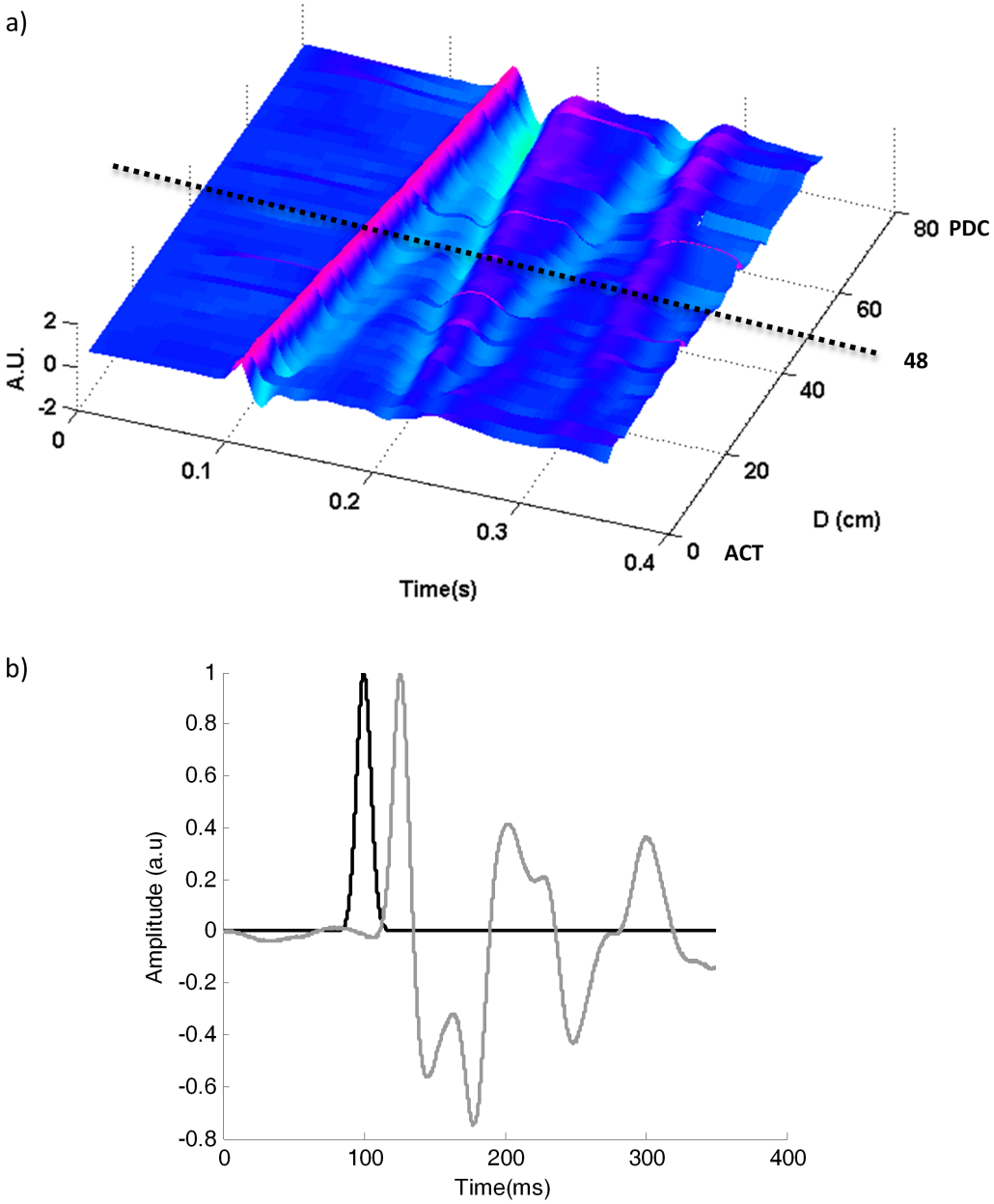

Figure 3.28: Test bench characterization using a Gaussian shaped excitation. (a) Shows the mesh map of the probe signal along the tube, (b) depicts the excitation (black line) and probe signal at the identified tube position (48 *cm* from the ACT) (grey line).

3.6 Conclusion

A multi-modular platform for non-invasive cardiovascular assessment has been designed and implemented. The APW module was the core of this thesis and it was described in detail. This module is based on a piezo disc sensor, analogue and digital signal electronic blocks able to recover

from the piezo signal the pressure waveform which can be displayed, and stored, sequentially.

Numerical simulations and test bench measurements, aimed at demonstrating the effectiveness of the proposed system, were carried out. The probe performed very well in a good number of fundamental accuracy tests using elastic tubes at the test bench. The maximum RMSE introduced by the electronic circuit itself is 0.33 % and the total average error (including the mechanical interface contribution) amounts to 1.80 %. The results lead to believe that the probe has the potential to be used as an alternative to costly devices in clinical trials.

3.6.1 Manufacturing Costs

The cost of a multichannel platform unit (considering the basic modules, APW and PWV) can reach 2160 €, at 2013 prices, as is presented in Table 3.5. There is a significant part in the cost of the system that can be saved by the development of a global control unit, replacing the NI 6210 (600 €). If considered the human resources costs for building and testing the prototype, as well as the required PC, the global system cost can reach 7000 €.

It is also important to note that the presented prices are unitary, meaning that buying in large quantities will significantly decrease the total price. Regarding commercial prices, equipment such as the SphygmoCor CPV or the Complior can reach 23500 € and 12500 €, respectively.

Table 3.5: Table of costs.

Module	**Manufacturing Costs (€)**
Box	520
Power supply	300
APW module	190
PWV module	60
NI 6210	600
ECG	140
Oximetry	200
Other components	150

All prices include VAT at 23%

4

Data Collection and Processing

This chapter presents information about collected data, acquisition protocols and processing tasks performed along this work. Firstly, the input variable groups are described, and then, the main risk subsets are identified (smoking history, lipid disorders, hypertension, diabetes mellitus and obesity). Processing techniques used along this work concerning pulse wave analysis are also introduced. The preliminary analysis includes baseline fluctuation elimination, pulse-by-pulse segmentation and pulse normalization. The delineator algorithm developed for the prominent points identification is introduced and, the selected morphological parameters are also detailed.

4.1 Recruitment of Participants

The data analysed throughout this thesis were collected in different groups of subjects. Some of the subjects were recruited through follow-up appointments of patients at the *Coimbra Hospital and University Centre* (*CHUC*), Portugal, while the others were obtained from the collaboration with volunteers at screening mobilizations. All of the subjects received detailed information about the study and agreed with criteria for participation. Figure 4.1 details the database cohorts fully described along this chapter. The grey rows report the experimental data used in the hardware and software tests. A review of demographic, clinical and non-invasive parameters included in the data analysis is presented in Table 4.1.

	Label		*Number of subjects*	*Year*
Bench Tests				2009-2010
First cohort	GroupF		20	2010
Control cohort	GroupC	Phase 1 Phase 2	134	2010-2012
Repeatability cohort	GroupR	Trial 1 Trial 2	32	2011-2012
Hypertensive cohort	GroupH		35	2011-2012
Angioplasty cohort	GroupA		6	2011-2012
Data prediction cohort	GroupD		10	2011-2012
	Valid subjects		**217**	

Figure 4.1: Database constitution by cohorts.

Table 4.1: The group variables.

Variable	Variable type	Description
Demographic	Patient related	Age, gender
	Procedure related	Date, time length, location, operator description
Clinical	Risk factor	Smoking habits, obesity
	Target value	Cardiovascular disease diagnosis
	Pressure values	SBP, DBP, HR
	Biochemical	TGL, CH, CH-LDL, CH-HDL, Aldosterone, Renin
	Other	Medication
Non-Invasive	Time and amplitude	SWTT, RWTT, LVET, SWA, RWA, DWA
	Ratios and Indexes	AIx, R1, R2, R3, R4, R5, R6
	Variance values	$RMSSD_{SPt}$, $RMSSD_{SPa}$, $RMSSD_{Pit}$, $RMSSD_{Pia}$, $RMSSD_{DWt}$, $RMSSD_{DWa}$, FWHM, RMSE

4.1.1 First Trials (GroupF)

This group was essential during the DAQP1 development. The data were used during the first tests to assess the system capability to deal with different waveform morphologies [166] and during the development of the algorithm dedicated to the APW prominent points identification [183].

4.1.2 Repeatability Cohort (GroupR)

This group was used in the validation tests, during two distinct trials, for the assessment of the differences between sessions (intra-subject variability), the left and right carotid differences, the monthly variability, and the operator dependency. The main characteristics of each dataset (Trial 1 and Trial 2) within this group are described in Table 4.2.

4.1.2.1 Trial 1

Dataset I consists of 12 healthy subjects (8 female, 4 male), under < 30 years. All of the subjects were non-smokers, and had no documented history of diabetes or cardiovascular disorders. All

measurements were made at the same time of day and at the same controlled room temperature ($22-23$ °Celsius) to minimize climatic variations. The subjects remained seated and quiet in a comfortable chair. Two successive monthly sessions, containing both left and right carotid data, were performed. At least 3 acquisitions of roughly $30-40$ *s* were made, *per* carotid, for each subject, *per* month. This dataset has a total of 2983 pulses, with $\approx$ $200-250$ pulses for each subject.

4.1.2.2 Trial 2

Dataset II consists of 20 healthy subjects (12 female, 8 male), also under $<$ 30 and with no documented history of cardiovascular disorders. The acquisitions were made by three operators with different levels of experience and training. Two operators (operator 1 and operator 2) have received one week of intensive instruction and training before recording measurements for the study. While the other (operator 3) had, at the time, more than 2 years of experience. Each operator performed several measurements, which consisted of three sessions for each one of carotid arteries (left and right), comprising a set of 18 acquisitions of about $60-90$ *s* for each subject. This dataset includes 11664 pulses, $\approx$ $500-550$ pulses *per* subject.

Table 4.2: GroupR characteristics for each one of the trials.

Characteristics	**Trial 1**	**Trial 2**
Subjects	12	20
Age range (years)	21-29	21-55
Total measurements per subject	12	18
Total number of pulses	2983	11664
Time acquisition per each measure (sec)	30-40	60-80
Repeatability-daily	$\surd$	$\surd$
Repeatability-monthly	$\surd$	X
Inter operator variability	X	$\surd$
Bilateral variability	$\surd$	$\surd$

4.1.3 Control Cohort (GroupC)

The acquisitions for this group started in 2010, denominated the Phase 1, with 20 subjects. However, the majority of data were obtained in 2012 (114 subjects), during the Phase 2 [184], in the collaboration with *Sudden Cardiac Death Screening of Risk Factors* study (*SCDSOS*) and during a mobilized screening at the Department of Physics of the University of Coimbra. This group consists of 134 healthy subjects, comprising 4471 pulses.

4.1.4 Hypertensive Cohort (GroupH)

This cohort includes 50 hypertensive subjects. Hypertension is diagnosed when SBP $\geq$ 140 $mmHg$ and/or DBP$\geq$ 90 $mmHg$, or if the patient was taking anti-hypertensive medication. The data were acquired during hospitalization, but prior to taking any medication. For each patient three sessions were performed during $30-40$ s.

4.1.5 Angioplasty Cohort (GroupA)

All the subjects at this group were monitored before and after a carotid intervention (due to a stenosis that partly blocked the artery blood flow). The intervention consisted in the angioplasty with stent placement, the "gold-standard" invasive procedures to treat a narrow or blocked carotid artery. The angiography procedure was used by injecting a contrast fluid to highlight the blood vessels, and making images of the highlighted vessels, such as described in Section 2.3.5

After local anaesthesia, a surgical cut is executed near to the groin and, a filter device is opened above the lesion for the stent implementation. The balloon is inflated at the stenosis plaque level. Finally, the filter device is kept open, as prevention, for an eventual embolic trapping. With this surgical technique, the blood flow is restored to the normal values, since the diameter of the vessel enlarges to the same one imposed by the stent. The APW was sequentially collected by the developed APW module and by the invasive equipment, Axiom Sensis XP (Siemens, Erlangen, Germany) [185]. A small segment of $3-4$ s for each subject was analysed.

4.1.6 Data Prediction Cohort (GroupD)

This group consists of 10 subjects that were used during the validation tests performed on the machine learning algorithms. The subjects were selected for both criteria: young subjects, and with waveform characteristics of individuals that are developing arterial disorders (such as, positive AIx values). To accomplish this selection, a preliminary waveform analysis was performed. This dataset comprises of 627 pulses.

4.2 Patient Characteristics

The clinical characteristics of this study population are shown in Table 4.3. The Kruskal-Wallis test was used to assess the differences amongst groups (see test description in Section 6.2.5). This test demonstrated significant differences in the analysis of the groups for all of the parameters, excluding height.

In GroupH, mean age is 58.89 ± 12.47 years, significantly higher than the observed for the other groups (excluding groupA). The same situation was verified for the BP values, that are higher comparing to the normal range (see Section 2.6.1) in the GroupA and in the GroupH. The values of 144.33 ± 40.08 $mmHg$ (SBP) and 81.00 ± 18.34 $mmHg$ (DBP) were verified in the GroupH and, the values of 147.37 ± 30.89 $mmHg$ (SBP) and 89.63 ± 16.29 $mmHg$ (DBP) were observed in the GroupA. It was also verified that the Body Mass Index (BMI) is superior to the ideal recommendations (< 25 kg/m^2) for the GroupH. In the GroupC, GroupR and GroupD, the BP ranges are within optimal values: SBP$<$ 120 $mmHg$, and DBP$<$ 80 $mmHg$. The mean age for the GroupD is 24.10 ± 2.81 years, similar to the value observed for the control group (GroupC) (21.90 ± 3.32) *years*, as expected, and the BP values are also within normal range.

Table 4.3: Popuplation characteristics for each analysed groups.

Variable	**GroupC**		**GroupR**		**GroupH**	**GroupA**	**GroupD**
	Phase1	Phase2	Trial1	Trial2			
Age	24.40 ± 4.06	21.90 ± 3.32	23.50 ± 2.43	23.35 ± 2.56	58.89 ± 12.47	72.5 ± 5.44	24.10 ± 2.81
Gender (M/F)	11/9	45/69	4/8	8/12	17/18	4/2	1/9
Smoker (Y/N)	2/18	10/104	0/12	5/15	5/30	M.D.*	0/10
Diabetes (Y/N)	0/20	0/114	0/12	0/20	4/31	M.D.*	0/10
Weight (kg)	66.75 ± 10.72	62.46 ± 10.24	59.66 ± 10.96	60.00 ± 8.70	75.24 ± 12.16	M.D.*	58.00 ± 5.85
Height(m)	1.71 ± 0.05	1.69 ± 0.09	1.66 ± 0.06	1.67 ± 0.08	1.65 ± 0.10	M.D.*	1.68 ± 0.07
BMI(kg/m)	22.65 ± 2.80	21.80 ± 2.59	21.45 ± 2.71	21.39 ± 2.09	27.82 ± 4.61	M.D.*	20.56 ± 1.43
SBP(mmHg)	110.50 ± 11.88	108.78 ± 11.59	105.25 ± 5.86	105.70 ± 14.06	147.37 ± 30.89	144.33 ± 40.08	102.10 ± 13.35
DBP(mmHg)	69.85 ± 10.47	69.52 ± 7.84	65.67 ± 6.71	69.95 ± 9.17	89.63 ± 16.29	81 ± 18.34	72.90 ± 11.35
HR(beats/min)	67.80 ± 11.02	70.11 ± 10.83	71.17 ± 10.64	70.45 ± 10.45	63.03 ± 6.76	72.5 ± 3.53	63.50 ± 8.89

M.D. - Missing Data

4.3 Risk Factor Characteristics

The factors used to describe the cases at risk were: the current or past history of smoking, the diagnosis of hypercholesterolaemia or hypertriglyceridaemis (hyperlipidemia), the presence of hypertension, the diagnosis of diabetes mellitus and the overweight. Figure 4.2 shows the distribution of the risk factors for all of the subjects in the database. The age distribution is depicted in Figure 4.3, where the predominance of subjects aged between 20 and 35 years is clearly evident.

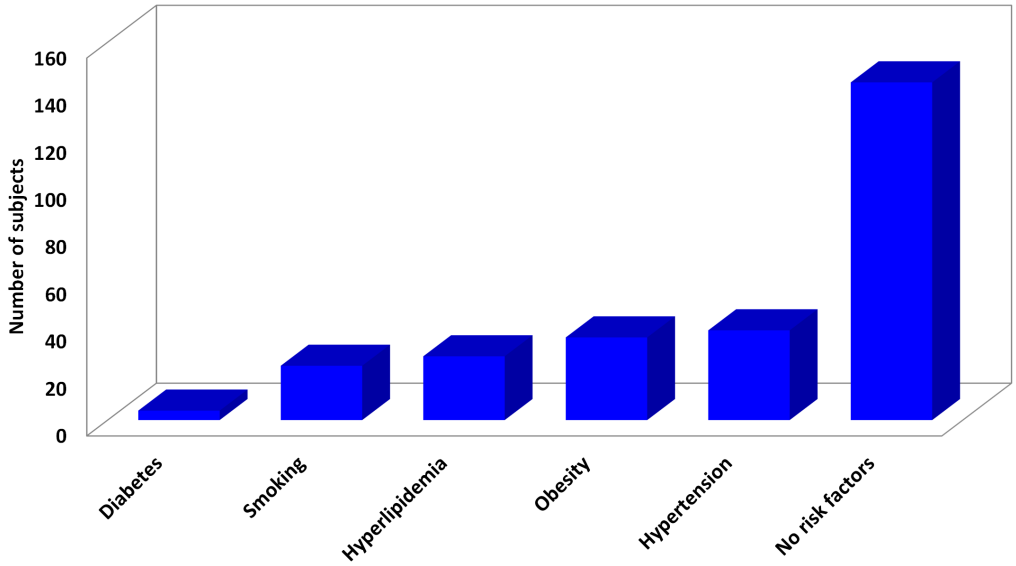

Figure 4.2: Risk factors distribution.

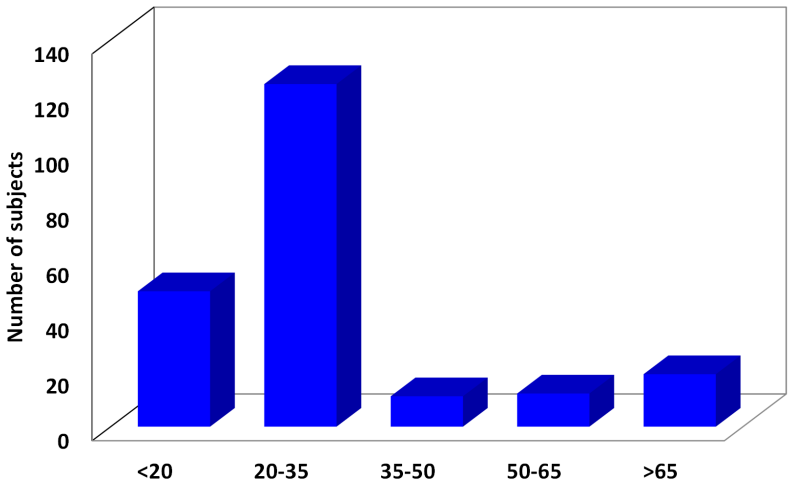

Figure 4.3: Distribution of subjects in study by age subsets.

4.4 Arterial Pressure Waveform Analysis

Pulse wave analysis relies on a clean waveform, in which the morphological features are reliably identified. The APW is the result of the the interaction of the amount of blood that is pumped out of the heart along the arterial tree. The APW contains important physiological information concealed on its morphology: the Systolic Peak (SP), the Point of Inflection (Pi) and the Dicrotic Wave (DW) [9]. SP results from the action of the left ventricle blood pumping, while Pi results from the both action, of the forward wave travelling along the arterial tree and of the backward wave returning towards the heart, from the reflection sites [2]. These waves superimpose originating a visible inflection change in APW profile. DW occurs when the aortic valve closes moving a small portion of the ejected blood back to the left ventricle. The prominent point locations are schematically represented in Figure 4.4, for a Type A (a) and a Type C (b)waveform. Depending of the AIx value the pulse wave type is defined as follows: when a negative value occurs the pulse is

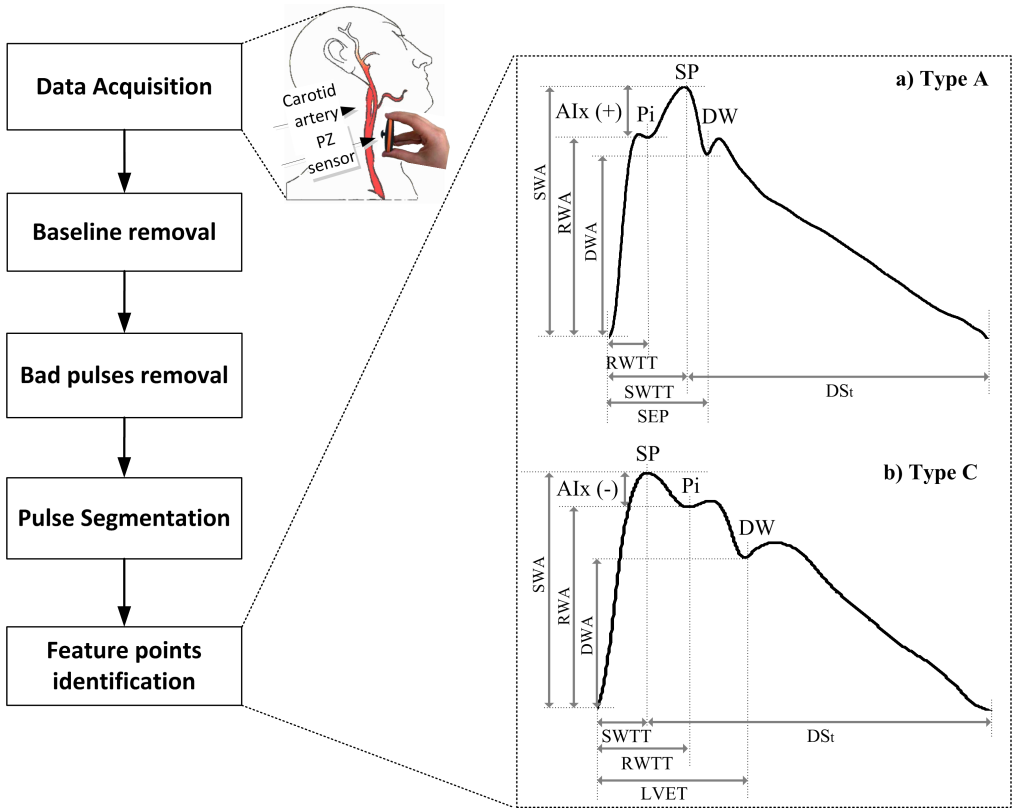

Figure 4.4: Schematic representation of the pulse wave analysis process, for a Type A (a) and a Type C (b) waveform, where the APW features are identified.

considered type C (characteristic of healthy subjects) and, when a positive value occurs the pulse is considered type A (characteristic of subjects suffering from arterial stiffness). Type B occurs for the intermediate cases, where the Pi occurs shortly before SP, and Type D when the Pi is not identified.

4.4.1 Pulse Waveform Modelling

Cardiac-like pressure waveforms were synthesized and used during the hardware and the software tests [181]. The set of pulses was synthesized using a weighted combination of exponentially shaped sub-pulses, each one representing a physiological component of the real cardiac pulses (the systolic stroke, the reflected wave, and the aortic reservoir). Each sub-pulse is built up by two consecutive exponential curves, describing the rising edge and another the falling edge. The general expression used is presented in Equation 4.1,

$$c(t) = \sum_{k=1}^{3} A_k^{e^{-\frac{t-D_{Rk}}{\tau_{Rk}}} - e^{-\frac{t-D_{Fk}}{\tau_{Fk}}}} \tag{4.1}$$

Prior to summing, the sub-pulses are submitted to a moving average filtering process in order to smooth the corners that, otherwise, would show up in $c(t)$. Table 4.4 describes the parameters used in Equation 4.1. The set of synthesized waveforms contains the interesting situations, namely the ones that occur during the AIx positive to negative transitions.

Table 4.4: Parameters used in Equation 4.1, where $k = 1$ corresponds to the systole component, $k = 2$ corresponds to the reflection wave and $k = 3$ corresponds to the dicrotic wave, and w was used to sweep the range of interesting Pi situations.

Parameter	**Description**	**Systole wave** $(k = 1)$	**Reflected wave** $(k = 2)$	**Dicrotic wave** $(k = 3)$
$A(k)$	Amplitude	1	0.2	0.3
$D_{Rk}(s)$	Delay of rising exponential	0.1	$0.1201 + (w \times 0.003)$	0.35
$\tau_{Rk}(s)$	Rising exponential constant	0.001	0.01	0.001
$D_{Fk}(s)$	Delay of falling exponential	0.192	$0.17 + (w \times 0.003)$	0.45
$\tau_{Fk}(s)$	Falling exponential constant	0.0007	0.0006	0.0001

4.4.2 Baseline Flutuations

The baseline flutuations can occur due to electrical signal fluctuations, small motion of the probe attached to the patient's neck or patient breathing. The baseline modulation can lead to a misleading and inaccurate determination of the characteristic points. A methodology to remove the baseline fluctuation from the raw signal by applying a baseline adjustment was employed, ensuring that the APW is precise and valid. Firstly, the baseline index points are accurately determined $y(n)$, such as depicted in Figure 4.5. These points match the onset and offset of APW pulses, as shown in Figure 4.5. Then, the baseline fit is linearly interpolated from the baseline index points, correcting the signal by vertically adjusting each sample.

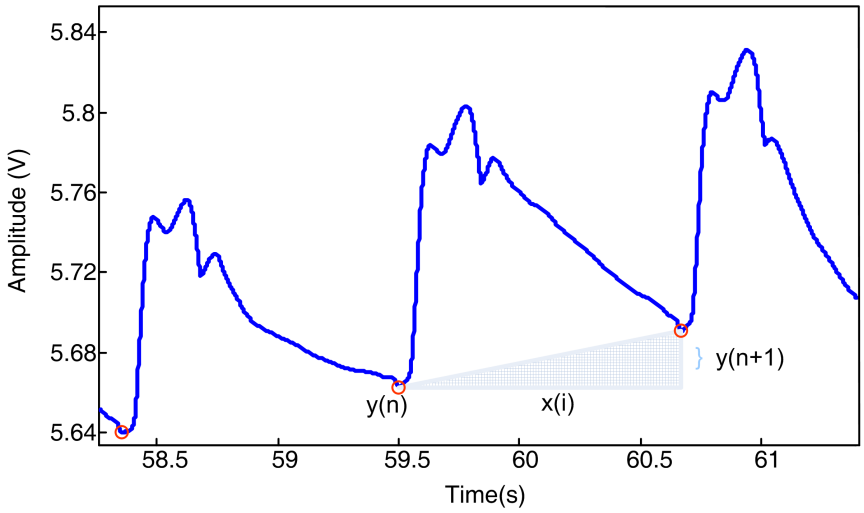

Figure 4.5: Method adopted for baseline correction, where $y(n)$ and $y(n+1)$ are adjacent APW onset pulses.

4.4.3 Criteria for Pulse Flagging

The morphological APW pulse quality is an important issue that must be taken into account. The artefacts, that can occur during data acquisition, are responsible for volume changes that compromise the quality of the signal. So, after the baseline removal, each pulse must be analysed and several morphological indexes computed by a series of constraints in order to check abnormal beats. The criteria were defined based on the normal range values for systolic and diastolic amplitudes, and pulse width. If any abnormality criterion is matched, the pulse is removed (see Table 4.5).

Table 4.5: APW features and abnormality criteria for pulse flagging.

Feature	Description	Abnormality Criteria
As	Systolic amplitude	$As > 3 \times Am$
Am	Mean systolic amplitude	$Am < \frac{As}{2}$
AD	Diastolic amplitude	$AD < 0$
T	Pulse Width	$T > 1.5 \times Tm$

Tm – Mean width of all pulses

4.4.4 Pulse Segmentation and Normalization

The APW signal segmentation into individual pulses and the amplitude normalization is an important process during data pre-processing. Figure 4.6 depicts the effect of baseline removal and pulse segmentation procedures. A raw time signal segment of 25 seconds, corrupted by baseline fluctuation, is exhibited in Figure 4.6(a) and the baseline removal result is shown in Figure 4.6(b). Pulse-by-pulse analysis is shown in Figure 4.6(c), where the variability among pulses is visible. In Figure 4.6(d) the prominent points (SP, Pi and DN) are pointed out.

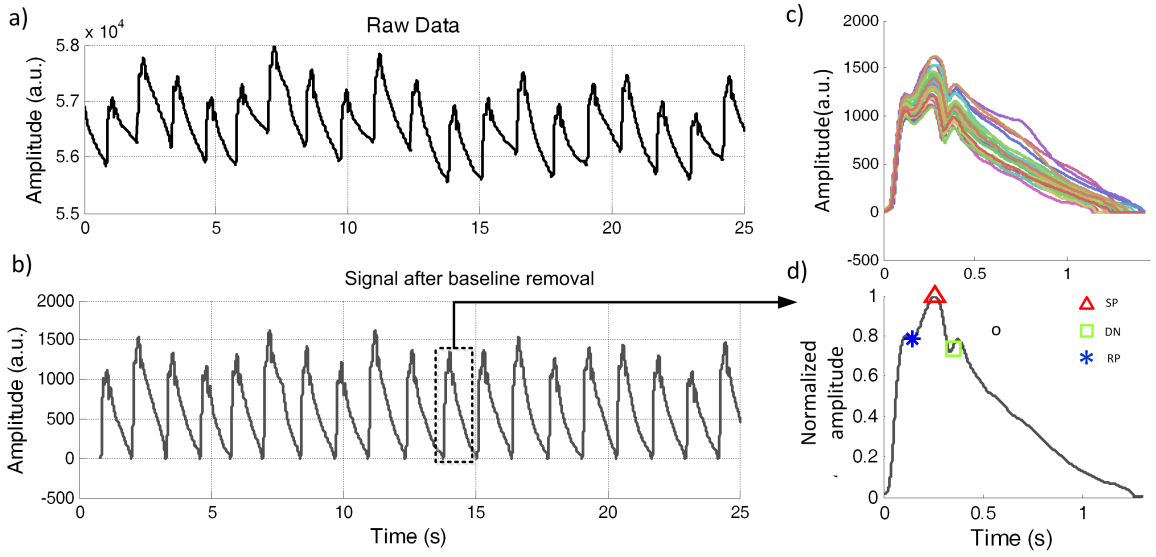

Figure 4.6: Pulse wave analysis, (a) raw data; (b) baseline removal; (c) pulse segmentation and (d) the prominent points marked in a pulse during the segmentation process.

4.4.5 Pulse Contour Algorithm Implementation

Several waveform delineators have been reported in literature concerning the APW characterization, such as the ones presented by Donelli *et al.* [186] and Oppenheim *et al.* [187]. De Melis *et al.* [188] also presented an interesting approach by the use of a discrete waveform transform (DWT), namely the ”mother” wavelet Daubechies 4 (Db4), in the determination of the temporal dynamic properties. The use of the second derivative of photopletysmographic (PPG) signals is also commonly referred as a reliable technique by other authors [189].

The first results were achieved during MSc thesis (2009) [190], by the implementation of algorithms based on the DWT and Probability Density Function (PDF) [181, 190]. Wavelet transform is a time-scale representation, based on the multi-resolution signal decomposition, which allows to follow the spectrum of the frequencies into the signal. The wavelet transform is an appropriate tool to rapidly locate changes in physiologic signals, such as in APW and ECG signals. PDF relies on the creation of a local maximum for the amplitudes close to a inflection [191]. Unfortunately, several maxima are also created whenever the signal amplitude is slow varying, as happens close to its peaks. To make things worse, these peaks can occur for amplitudes of the same order of magnitude of the Pi point, making the identification of this point a task very hard to accomplish. The preliminary results, using synthesized waves, have demonstrated the superiority of the wavelet approach.

The purpose along the current work was the development of a robust set of algorithms capable of dealing with a wider variability of physiological conditions. The minimal requirements remain on the capability to perform the identification in the waveform (in time and relative amplitude) of the most relevant points of the cardiac cycle and make them available for further analysis in an understandable and clear way. An algorithm based on the first order derivative was implemented, as described in Figure 4.7. The use of the first derivative curve justifies that, immediately following the data acquisition, a low pass filter needs to be implemented to suppress high frequency noise, that would otherwise make the first derivative signal unusable. The 30 Hz cutoff frequency of the filter ensures that the interesting range of frequencies is preserved. A short overview about each of the identified points is presented as follows:

- **Systolic Peak**

  The SP identification is carried out using a local extreme identification routine. A maximum is detected, only in case the value of the previous minimum differs by a configurable value (referred to as delta). As, the cardiac cycle shows a variable number of local maxima (2 to 4 depending of the wave type and of the presented artefacts), this routine must run repeatedly (for different values of delta) to ensure the correct SP identification.

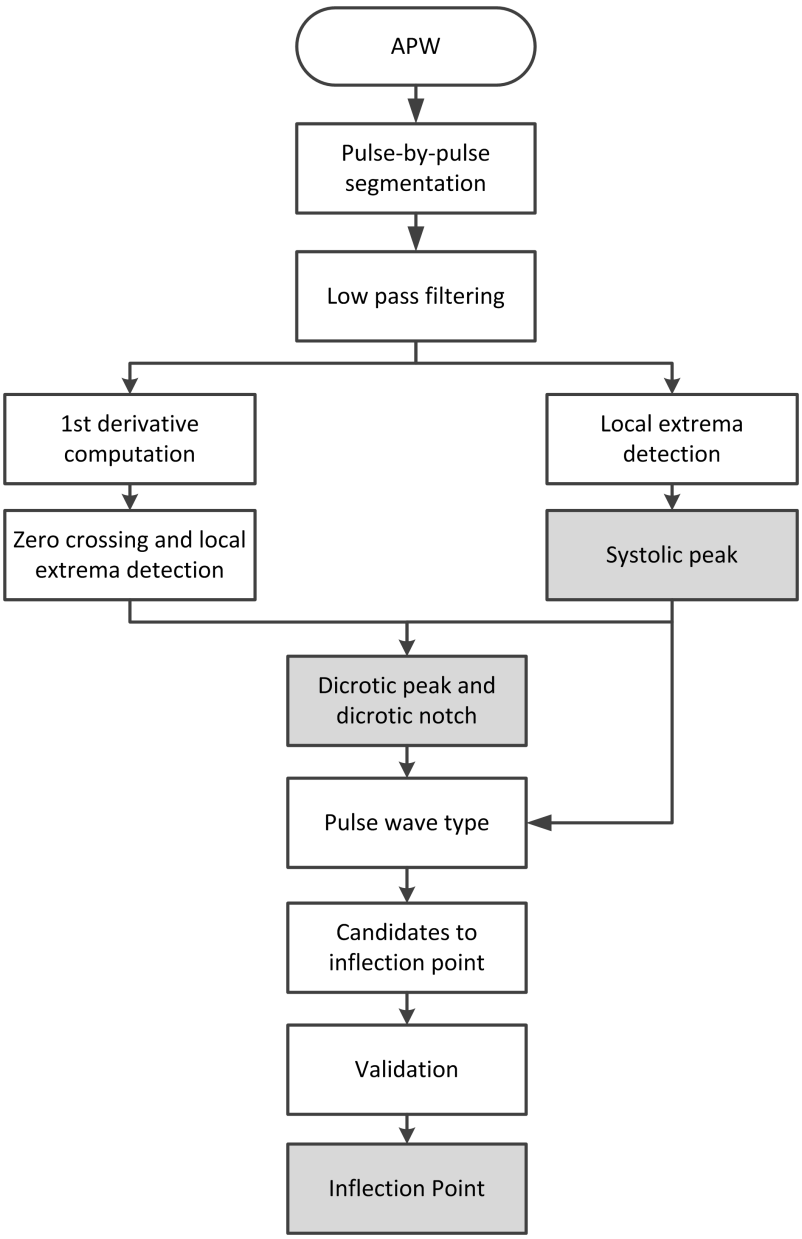

Figure 4.7: Flowchart representing the algorithm implementation.

- **Dicrotic Notch**

  The SP identification combined with zero crossing values of the first order derivative is used in the DW features identification, notch (DN) and peak (DP). The major problem are the oscillations that inevitably may occur, and mask the true DN and DP.

- **Inflection Point**

  The Pi identification is carried out in three steps (validation procedure): the localization of candidates, the elimination of oscillations and the comparison with APW maximum.

4.4.5.1 Performance Evaluation

Some acquisitions were carried out in a small universe of eight volunteers. In a first step, the prominent points (SP, DN, DP and Pi) were identified using the algorithm under test, and then, a Human Expert Engineer (HEE) carefully inspected each APW pulse and manually annotated the same points. Both annotations were compared and classified in three classic types: the true positives (TPs) and the true negative (TNs), and for the discrepant ones, the false negatives (FNs) and the false positives (FPs). The HEE results were taken as reference, and a 8 *ms* threshold value was adopted, such as indicated by others authors [192, 193]. An identical procedure was carried out for the amplitude parameters, using a 1 % of the maximum pulse amplitude as the threshold value. Sensitivity (S), positive predictive value ($P+$) and *error* associated to the algorithm were computed, according to the following equations:

$$S = \frac{TP}{TP+FN} \tag{4.2}$$

$$P^{+} = \frac{TN}{TN+FP} \tag{4.3}$$

$$Error = \frac{TP+TN}{TP+TN+FP+FN} \tag{4.4}$$

Tables 4.6 and 4.7 show the sensitivity, the positive predictive value and the error verified during the time and amplitude analysis, respectively. As expected, Pi shows the worst performance,

Table 4.6: Validation of algorithm performance (time) in comparison to HEE annotations.

Feature	Pulses	TP	FP	FN	Error (%)	S (%)	P^+(%)
SP	173	171	2	0	1.15	100	98.84
DN	173	165	8	0	4.62	100	95.37
DP	173	169	4	0	2.31	100	97.69
Pi	173	159	8	6	8.38	96.36	95.20

Table 4.7: Validation of algorithm performance (amplitude) in comparison to HEE annotations.

Feature	Pulses	TP	FP	FN	Error (%)	S (%)	P^+(%)
SP	173	173	0	0	0	100	100
DN	173	171	2	0	1.15	100	98.84
DP	173	172	1	0	0.58	100	99.42
Pi	173	158	9	6	8.98	96.34	94.61

due to the superior difficulty associated to its identification, observed by the HEE well as by the implemented algorithm. Relatively to time measurements, average values of 4.20 %, 99.09 % and 96.77 % for *error*, *S* and *P*+ were observed, respectively. And, for the amplitude analysis were observed values of 2.68 %, 99.08 % and 98.22 % for the same parameters.

Figure 4.8(a) shows the correlation plots between the HEE and the algorithmic values (SP and Pi analysis), as well as the the corresponding Bland-Altman plots. The straight line fittings for the SP and Pi time analysis show an excellent correlation, where R^2 is the square of the sample correlation coefficient, between both estimates ($R^2 = 0.996$ and $R^2 = 0.907$, respectively). The developed algorithm, in average, overestimates the SP by 0.009 *ms* and underestimates the Pi by 0.99 *ms*. To understand the 100 fold factor between the estimation errors between SP and Pi, one should bear in mind that, in the APW curve, SP is a peak and Pi is an inflection. This makes the SP estimation much easier for the HEE, as well as for the algorithm.

It was also observed that the amplitude errors summarized in Table 4.7 are much lower than the corresponding time errors (Table 4.6). In fact, as all points are associated to a peak or an inflection, the first derivative of the APW curve shows close-to-zero values in their vicinity, hence the small amplitude estimation error. This fact is visible in Figure 4.8(b) that shows the correlation observed for Pi amplitude values, as well the corresponding Bland-Altman plot.

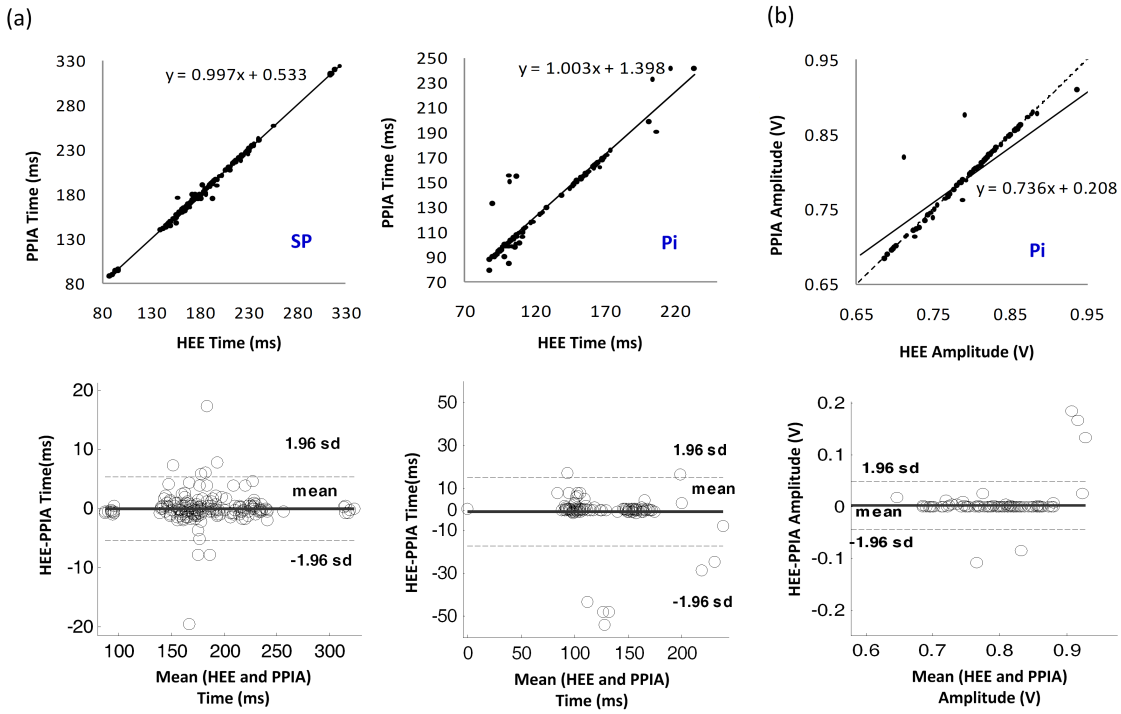

Figure 4.8: Agreement between the SP and Pi obtained via PPIA and HEE, measured using regression plots (top) and Bland-Altman plots (bottom), (a) time information of SP and Pi, (b) relative amplitude of Pi.

4.4.6 Pulse Waveform Features

Most of APW analysis frameworks individually process each hemodynamic parameter and do not evaluate the dependencies among features. The proposed approach allows for the design of new hemodynamic morphology vectors and techniques for multiple APW analysis. This approach improves the arterial pulse understanding, especially when compared to the traditional single-parameter analysis, where the failure in one parameter measurement component, such as Pi, can jeopardize the whole evaluation. Several hemodynamic features were explored and are described as follows:

- **Time/amplitude Intervals**

  The most important APW intervals are considered for each pulse. These parameters include time intervals, the time to the systolic peak (SWTT), the time to the point of inflection (RWTT) the time to the dicrotic wave (LVET) and the downstroke time (DS_t; and amplitude considerations, the height at the systolic peak (SWA), the height at the inflection point (RWA) and the height at the dicrotic wave (DWA).

- **Ratios and Indices**

  Although time and amplitude parameters are effective, they lack a complex analysis. Other parameters and indices were studied based on time and amplitude intervals, as presented in Table 4.8.

Table 4.8: Ratios and indices used in the multi-parametric approach.

Attribute	Description	Units
$R1 = \frac{SWTT}{T}$	Upstroke time ratio	[-]
$R2 = \frac{SWTT - LVET}{T}$	Downstroke time ratio calculated between the systolic and the diastolic waves	[-]
$R3 = \frac{DS_t}{SWTT}$	Time ratio calculated between the global downstroke time (DS_t) and the upstroke time $(SWTT)$	[-]
$R4 = \frac{DWA}{SWA}$	Amplitude ratio between the dicrotic and the systolic heights (SWA)	[-]
$R5 = \lvert SWA - RWA \rvert$	Amplitude difference module between the systolic and the inflection points	[V]
$R6 = \frac{RWA}{SWA}$	The systolic and the inflection points amplitude ratio	[-]

The well-established AIx, which measures the augmented pressure is computed according to Equation 4.5, that was defined by Murgo *et al.* [182], where A is the normalized amplitude.

$$AIx(\%) = \pm \frac{SWA - RWA}{A} \times 100 \tag{4.5}$$

The key issue concerning the AIx computation is the Pi identification, that can occur after, or before, systolic wave front. Negative values are used to distinguish pulses where Pi arrives after the systole ($RWA - SWA$), from the positive values that occur when the Pi occurs before the systolic peak ($SWA - RWA$).

- **Pulse Variance**

  Some statistical measurements were performed to assess the variance associated to the prominent points locations, such as presented in Table 4.9. The Root Mean Square Successive Difference (RMSSD) was computed for all of the variables, and the overall list includes: RMSSD_{SPt}, RMSSD_{SPa}, RMSSD_{Pit}, RMSSD_{Pia}, RMSSD_{DWt} and RMSSD_{DWa}. Additionally, the pulse morphology variability was assessed using the RMSE measure, taking as reference the average pulse. Finally, the Full Width at Half Maximum (FWHM) was also evaluated.

Table 4.9: Pulse variance attributes.

Pulse Variance	**Description**	**Units**
RMSSD_x $\sqrt{\frac{\sum_{i=1}^{n-1}(x_{i+1}-x_i)^2}{n-1}}$	RMSSD, where x is: – SP_t (represents SWTT) – SP_a (represents SWA) – DW_t (represents LVET) – DW_a (represents DWA) – Pi_t (represents RWTT) – Pi_a (represents RWA)	[%]
RMSE $\sqrt{\frac{\sum_{i=1}^{n-1}(x_{1,i}-x_{2,1})^2}{n}}$	RMSE (measured in reference to the average pulse)	[%]
FWHM	Full Width at Half Maximum	[s]

5

Clinical Assessment Support Systems

The aim of this chapter is the exploration of data mining techniques to deal with the vectorized features extracted from the APW. Different machine learning algorithms were used: Random Forest and J48 (decision trees), BayesNet (probabilistic), RIPPER (rule-based induction) and Artificial Neural Networks (ANNs) topologies. Additionally, clustering techniques were also implemented. The performance evaluation included the comparison with traditional parameters: AIx, lipidic measurements (e.g., cholesterol, cholesterol-LDL, cholesterol-HDL and tryglicerides), as well as risk scores commonly used in clinical practice.

5.1 Motivation

The major motivation of this chapter was the development of an efficient waveform analysis methodology able to deal with multiple APW parameters. There is growing interest in data mining methodologies that can ease the understanding of large biomedical datasets. The application of machine learning algorithms can ease the data analysis, ensuring the learning of the system from the data at which it is exposed to. Once trained, the system is able to predict a result, for a set of new inputs based on the internal rules previously established. This methodology could bring an interesting impact comparatively to the traditional cardiovascular tools used for risk prediction.

The Weka (Waikato Environment for Knowledge Analysis), a free machine learning platform, is a very powerful tool that can be used for the analysis of the predictive value of each attribute, and for the implementation of classification and clustering techniques. Additionally, it is a very efficient, versatile and affordable tool, A detailed description is presented in Appendix C.

There are several terms that need to be understood, before delving into the question of how machine learning schemes operate. In this work three main learning schemes were explored: classification, clustering and prediction. Classification is the process of finding a model to describe and to distinguish data classes. Typically, the derived model can be represented, *e.g.*, in the form of classification rules, decision trees or neural networks.The purpose is to build a model able to be used in the prediction of the class of objects whose class label is unknown. Prediction can be viewed as the construction and use of a model to assess the class of an unlabelled object, or to assess the value or value ranges of an attribute that a given object is more likely to have. Finally, clustering methodology can be used to analyse data objects without class label. The objects are clustered or grouped based on the principle of maximizing the intraclass similarity and minimizing the interclass similarity [158].

The input to a machine learning scheme is a set of instances. These instances are the things that are to be classified, associated or clustered. Each instance is characterized by the values of attributes that measure different aspects of the instance. There are many different types of attributes, although typical data mining schemes deal only with numeric and nominal. There is a broad distinction between quantities that are numeric and ones that are nominal. Numeric attributes, sometimes called

continuous attributes, measure numbers, while nominal attributes take on values in a pre-specified, finite set of possibilities and are sometimes called categorical [194].

5.2 Feature Subset Selection

Attribute selection is an important task that allows the improvement of dataset analysis, in a process where the unneeded information is identified due to its low predictive value. An attribute may not be selected, when presents too many missing values, and, there is (almost) no difference between the values. The identification of redundant and/or irrelevant attributes before the training process allows the optimization of the memory space and ensures time machine computing speed requirements. There are several algorithms that can be used for this purpose, such as the *InfoGainAttributeEval* (available on WEKA) that is a single-attribute evaluator algorithm very suitable for the analysis. It uses a Ranker search method that evaluates the selected attributes, by measuring their information gain, according to the defined classes using a discretization method [194]. So, the attributes are ranked according to their average merit (expressed as a number between 0 and 1).

5.3 Classification

Several experiments were made in order to select the most suitable algorithms both from the standpoint of discrimination ability, and the time spent to make a decision. The evaluation was focused on different types of machine learning algorithms, namely: decision trees, association rules, Bayesian networks, and Artificial Neural Networks (ANNs).

5.3.1 Single Classifiers

5.3.1.1 J48

The J48 is the Weka implementation of C4.5 algorithm, the most popular tree classifier, which was developed by Quinlan (1993) [195]. Its implementation is based on a non-backtracking approach. J48 uses two heuristic criteria to rank possible tests: the information gain and, the default gain

ratio. After the building process, each attribute test, along the path (from the root to the leaf) becomes a rule antecedence (precondition), and the classification at the leaf node becomes the rule consequence [158]. Decision trees are easy to interpret, capable to work with missing values, categorical and continuous data, characteristics that appeal for its application in medical field.

5.3.1.2 Random Forest

Random Forest is a meta-learner algorithm comprising many individual trees, which has been developed to operate quickly over large datasets [196]. Each tree depends on the values of a random vector independently sampled and with the same distribution for all trees in the forest. For forests, the generalization error converges to a limit as the number of trees in the forest becomes larger [197]. The main advantages are related to the robustness to noise, and the fast computation ability. However, the output can be difficult to interpret which is a disadvantage in medical environment.

5.3.1.3 RIPPER

RIPPER is a rule-based induction algorithm. After producing a rule set for the class, each rule is reconsidered using reduced-error pruning, before proceeding to generate rules for the next class. RIPPER was evaluated through JRIP which is the implementation of RIPPER in Weka [194]. The main advantages of using this algorithm are: a clear set of classification rules used, and a high speed computation.

5.3.1.4 BayesNet

BayesNet is a Bayesian network based on probabilistic graphical models. It was developed to deal with nominal values (in case of numeric values, they are pre-discretized) and absence of missing attributes [198]. Bayesian networks have a significant impact on the modelling and in the analysis of the patient data. The major advantages are: the easy interpretation of the results and the robustness in dealing with missing data.

5.3.1.5 Artificial Neural Networks

ANNs are sophisticated analytical techniques, inspired by the biological nervous system. A network is a set of connected input and output units, in which each connection has a specific weight associated to it. The network learns by adjusting the weights. Even though ANNs need high computational resources and rigorous parameter tuning (*e.g.*, the number of hidden neurons), being sometimes difficult to interpret, they present advantages that make their effectiveness undeniable. Amongst the several set of architectures, it should be stressed out, the outstanding accuracy, the ability to work with continuous attributes and the high tolerance to redundant attributes.

The most popular ANN architecture is the Multi-Layer Perceptron (MLP), which consists of an input unit, an output layer and one or more hidden layers that can be determined by the user. It uses a backpropagation learning algorithm. The network inputs pass through the input layer, and then, they are weighted and fed to a second layer (the hidden layer). The outputs of the hidden layer can be inputs to another hidden layer or, inputs to the output layer, which emits the network's prediction for the given tuples.

The Radial Basis Function (RBF) network is another popular approach. They are typically configured with a single hidden layer of units whose output function is selected from a class of functions called basis functions [199]. Being similar to the backpropagation in many aspects, radial basis function networks have several advantages. They are usually trained much faster than backpropagation networks. They are also less susceptible to problems with non-stationary inputs.

5.3.2 Classifier Ensemble - Bagging and Voting

The combination of multiple classification algorithms is based on the idea that each learning technique has its limitations and that no single technique can significantly outperform the others [200]. Essentially, there are two general groups of classifier fusion techniques, homogeneous ensemble and heterogeneous ensemble.

Homogeneous ensemble uses the same learning technique on different distribution of training datasets. Bagging is a well-known homogeneous ensemble technique that generates multiple sample

datasets by repeatedly sampling with replacements from training datasets. Each of them can be used to build a prediction model using the same learning technique. The output is estimated by averaging all the predictions [201].

Heterogeneous ensemble operates mainly on the classifier outputs combining different supervised learning techniques. Each classifier is first build on training data and then associated with a weight and measured by its prediction accuracy [201]. Then, a specific combination rule is selected for the final estimation of the output value. The voting system is one of the strategies that can be used in the combination of the results from the different classifiers.

5.4 Clustering

Clustering is defined as an unsupervised learning technique that can be used in the analysis of the independent variables (all variables except class). In fact, the capability to deal only with independent variables is the main feature that distinguishes clustering from classification techniques. For this reason, clustering techniques may be best used for exploratory studies, especially those that encompass large amount of data. During the cluster assignments, the pulses are independently grouped according to their cluster similarities. The maximum level of similarity for each subject occurs when all pulses fit in only one cluster. Expected Maximization (EM) and k-means clustering algorithms are the most studied.

5.4.1 k-Means

The k-means algorithm is a type of partitional clustering that continuously iterates until a specific criterion function (usually the square error) converges. It acknowledges the number of desired cluster inputs (k) and divides the set of objects (n) into k clusters. The result is a higher intra-cluster and lower inter-cluster similarity. The cluster similarity is measured considering the mean value of the objects contained in the cluster, which is, in fact, the cluster centroid. k-Means clustering is relatively scalable and efficient in processing large datasets. However, it cannot handle categorical attributes and it is not suitable for dealing with non-convex shapes. Also, it is quite

sensitive to the presence of noise and outliers. It requires an efficient data pre-processing before its application [158].

5.4.2 Expectation-Maximization

The Expectation-Maximization (EM) clustering algorithm is a complex probabilistic extension of the *k*-means method that primarily differs by the way how the initial groups are obtained. Instead of assigning each object to a cluster, with which it is most similar, EM assigns each object to a cluster according to a weight that represents the probability of membership. In this manner, there are no strict boundaries between clusters, and new means are determined based on weighted measures [158].

5.5 Risk Scores - Reference Method

Such as described in Section 2.7.1 there are several risk scores that have been developed to assist clinicians in the CVD risk assessment. Four risk scores were considered along this work namely, FRS, PROCAM, ASSIGN and SCORE. The FRS was developed from a general population in Framingham, Massachusetts, USA, while the remaining scores derive from European studies. The largest is the SCORE, since it consists of $\approx$ 205 000 subjects from 12 cohort studies from European countries. The other two, ASSIGN and PROCAM, were developed in Scotland and Germany, respectively. In general, the tools have distinct characteristics, different risk factors, specific disease, event type, period of time and cohort locations. Table 5.1 summarizes the main characteristics of them. The evaluation platforms are shown in Appendix D.

These tools are important to help physicians in their daily practice. However, its application in different population remains a topic concerning attention. Estimation of cardiovascular risk can be constrained by some inaccuracies due to the heterogeneity of populations in study. There are several issues raised considering its application in different geographical regions according to the original areas of the study. The clinical practice demonstrated that missing values may be problematic and the majority of the current used risk scores are not prepared to deal with them.

Table 5.1: Risk assessment tools (10 years term).

Model	Patients	Country of Origin	Risk factors
FRS	8491	USA	Age, Gender, Total-CH, CH-HDL, SBP, BPT, SMK
SCORE	205178	Finland, Russia, Norway, Denmark, UK (England), UK (Scotland), Sweden, Belgium, Germany, Italy, France and Spain	Age, Gender, Total-CH, CH-HDL, SBP, and SMK
ASSIGN	13297	Scotland	Age, Gender, FH, DB, SMK, SBP, Total-CH and CH-HDL
PROCAM	5389	Germany	Age, CH-LDL, SMK, CH-HDL, SBP, PE, DB, and TGL

Total-CH − Total Cholesterol, BPT − Blood Pressure Treatment, FH − Family History, TGL − Triglycerides, SMK − Smoking habits, DB − Diabetes, PE − Previous Event, SBP − Systolic Blood Pressure, CH-HDL − High-Density Lipoprotein, CH-LDL − Low Density Lipoprotein.

The trend for overestimation the risk factor in low-risk populations and underestimation in high-risk groups has been successfully demonstrated by Cooney, *et al.* [149]. It is known that an examination of 5 % SCORE can equate to a $10-25$ % FRS risk, depending on which of the several FRS functions is selected [40]. Haq, *et al.* [202] studied several methods for risk estimation (FRS, PROCAM, Dundee, and British regional heart-BRHS) and the results demonstrated a close agreement between all these, regarding average risk and showed moderate agreement for estimation among individuals. Finally, it was also concluded that FRS function is acceptably accurate in northern European populations.

5.6 Methods

5.6.1 Training

Commonly, there are three types of datasets: the training data, the validation data and the test data. The training data are used by the learning schemes to come up with classifiers, while the validation data could be used in the parameters optimization, or in the selection of a particular

one. The test dataset is commonly used in the computation of the error rate. All these datasets must be independently chosen. For a larger database this process is easier, since a larger and more representative sample can be selected for training while another independent large sample can be used for training purposes. For these cases, the error rate of the test set is a more reliable measure. However, whenever a larger database is unavailable and, the amount of data is limited, the best solution is reserve some data for testing, and use the remaining data for training, thus assuring the representatively of samples. Cross-validation is a simple form to ensure this. In this process, a fixed number of folds (or partitions) is defined and then, each of the folds is used once for testing, while the remaining folds are used for training. The procedure is repeated, so that, at the end, every instance has been used once in the testing procedure. Extensive tests on numerous different datasets and learning techniques, have shown that 10 is a good number of folds to be used in the error estimation [194]. In this case, the original dataset is partitioned into 10 subsets (folds) of equal size, $P1, (...), P10$. The global accuracy is the average of the accuracy of the 10 trials.

5.6.2 Classifiers Performance

Data from GroupC and GroupH were used, during the training procedure of the classifiers. Several experiments were made in order to select the most suitable methods both from the standpoint of discrimination ability, as well as, from the time spent to make a decision [158].

Such as depicted in Table 5.2, the measurements are generally presented in a confusion matrix. The True Positive (TP) and True Negative (TN) values represent the correct classifications; the False Positive (FP) values occur when the outcome is incorrectly predicted as positive and; the False Negative (FN) values occur when the outcome is incorrectly predicted as negative. Each column of the matrix represents the predicted classes, while each row represents the actual classes.

These values can be used in the computation of several indices (error rate, sensitivity, specificity

Table 5.2: Confusion matrix of observed class *vs.* predicted class.

	Positive (predicted)	**Negative (predicted)**
Positive (actual)	TP	FP
Negative (actual)	FN	TN

Table 5.3: Parameters used in the assessment of the classifiers' performance.

Parameter		Description
Sensitivity	$\frac{TP}{TP+FN}$	Percentage of positive labelled instances (actual condition) that were predicted as positive
Specificity	$\frac{TN}{TN+FP}$	Percentage of negative labelled instances (actual condition) that were predicted as negative
Positive Predictive Value (Precision)	$\frac{TP}{TP+FP}$	Percentage of correct positive predictions
Negative Predictive Value	$\frac{TN}{FN+TN}$	Percentage of correct negative predictions
Error Rate	$\frac{TP+TN}{TP+TN+FP+FN}$	Percentage of correct predictions

and precision) such as is shown in Table 5.3. It is also important to consider the training time along experiments. The following rank was adopted: * $[0-2\ s]$, ** $[2-10\ s]$, *** $[10-20\ s]$, **** $[20-60\ s]$, ***** $[>60\ s]$.

On the other hand, the Receiver Operating Characteristic (ROC) curve is a graphical plot of the sensitivity, given by the true and the false positive rate, being a good tool for the assessment of the classifiers' performance. The Area Under Curve (AUC) of the ROC is often used for the models comparison. A larger AUC value corresponds to a more accurate classifier (ideally $AUC = 1$).

5.6.2.1 Learning Curves

The study of the "learning curves" is an interesting point to consider during the classifiers evaluation, giving information about the effect of gradually increasing the amount of training data [203]. This process allows to gain insight into the generalization of the capabilities of each classifier. The

classifiers can be trained with a varying number of subjects, randomly selected from the whole dataset along repeated trials. As the number of subjects increases, it is expectable that the training performance should approach an asymptotic value, for each classifier.

5.6.2.2 Artificial Neural Network Performance

The definition of the best ANN configurations, especially for small datasets, is a complex and challenging task. An extensive search, even though it could lead to an optimal set of features, is impractical in terms of effort and computational time. Along this work different parameter configurations were tested and, the accuracy, the sensitivity, the specificity and the precision values comparatively analysed.

5.6.3 Predictive Models

The validation of the models was carried out using GroupD (see Section 4.1). The models obtained from the training of J48, Random Forest, BayesNet and JRIP were loaded to the class prediction of each pulse, individually. The majority-voting system (see Section 5.3.2) was used to merge the labels from each of the considered models.

The following nomenclature for predicted class labels was adopted: the labels correlated to the APW patterns from GroupH were coded as A, while the ”normal” APW patterns similar to those of GroupC (negative AIx values) were coded as B. Each subject can have pulse instances observed in both classes. In this case, the final value is presented as a percentage of the class A and the class B, within the set of analysed pulses.

5.6.4 Clustering

Clustering methods are very useful when little or no information is known about the dataset, and a primary visual data inspection is essential. GroupC and GroupH were selected and analysed for the set of characteristics available for each one. Different nomenclatures were adopted to help understanding the data at hand, numeric labels (1, 2, ...) for the GroupC and alphabetic labels (A, B, ...) for the GroupH.

During the cluster assignments for each subject, the pulses were independently grouped according to their cluster similarities. The maximum level of similarity within a cluster occurs when all pulses for each subject fit in only one cluster group. Therefore, the cluster risk is assessed as a percentage of the pulses in the fit of the predefined cluster at risk.

5.6.4.1 Number of Clusters

The choice for the number of clusters is often a difficult task. The clusters are evaluated according to the total squared distance criterion. The goal is to minimize the square-error that is the sum of the Euclidean distances between each pattern and its cluster centre.

The adopted strategy for the determination of the ideal number of clusters consisted, firstly, in the selection of two clusters ($k = 2$) followed by a tentative to split each of these clusters. During the process, clusters can also be merged if they are sufficiently close or, if there is too many patterns and an unusually large variance. As, the presence of outliers influences the number of clusters, these points should be eliminated from the data [15].

5.6.5 Risk Evaluation

The CHD risk was computed for each patient of GroupH based on the FRS, SCORE, ASSIGN and PROCAM functions. Firstly, the significance of correlation between the risk values and the other parameters at test was studied. The comparison included parameters from pulse wave analysis (APW parameters), BP values, and lipidic measurements.

The analysis was performed using statistic methodologies, to evaluate the normality of the distribution (Kolmogorov-Smirov test) and the correlations among variables (Spearman's rank correlation or Pearson product moment correlation). A full description of these methods is presented in Sections 6.2.2 and 6.2.5, respectively.

5.7 Results

5.7.1 Ranked Feature List

The *InfoGainAttributeEval* ranker method was used in the analysis of the set of features at test. Results are presented in Figure 5.1. The parameters are ranked according to the degree of correlation within the class, and the redundancy among them. Values near to 1 are obtained for those that are highly correlated with the class analysed and that show low correlation inter-parameters. The RWTT and SWTT time positions revealed to be the best predictors, as well as some of the computed ratios (R6, R1, R3 and R2) and index (R5). This task was essential for the discrimination of the most interesting features for further analysis. The parameters with low merit (a threshold of 0.1 was considered) were discarded. The overall list of parameters selected includes: RWTT, R6, R5, SWTT, R1, R2, R3, DWA, R4, LVET, AIx and RMSSD_{Pit}.

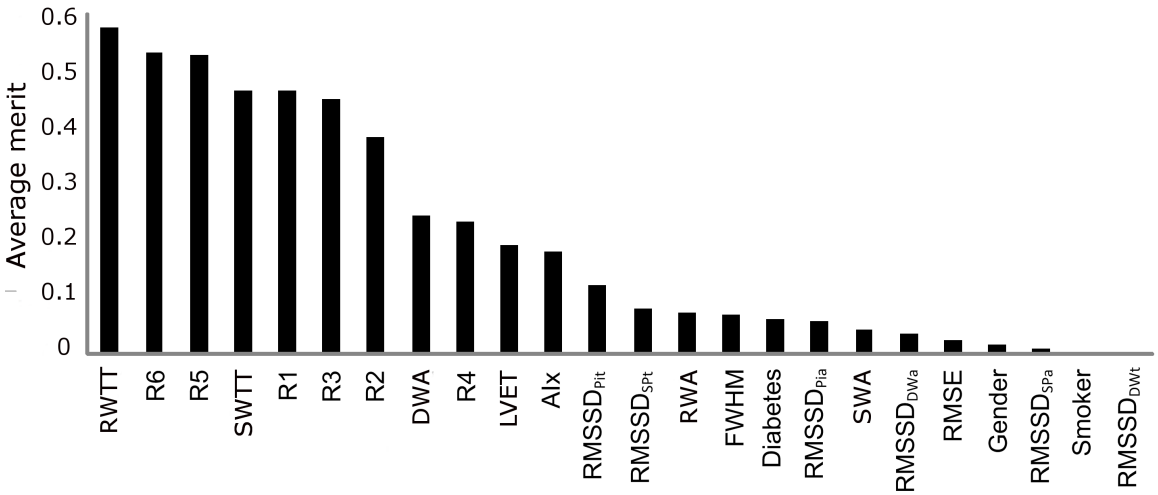

Figure 5.1: Ranking of analysis features based on the average merit.

5.7.2 Classifiers Configuration

5.7.2.1 Learning Curves

In Figure 5.2 are displayed the learning curves obtained varying the number of subjects in the dataset. The performance was assessed considering 5 trials for each configuration. The performance of all

the classifiers converged for a stable value, as presented in Figure 5.3. The Bayesian algorithms evidence the worst accuracy values comparatively to the other classifiers.

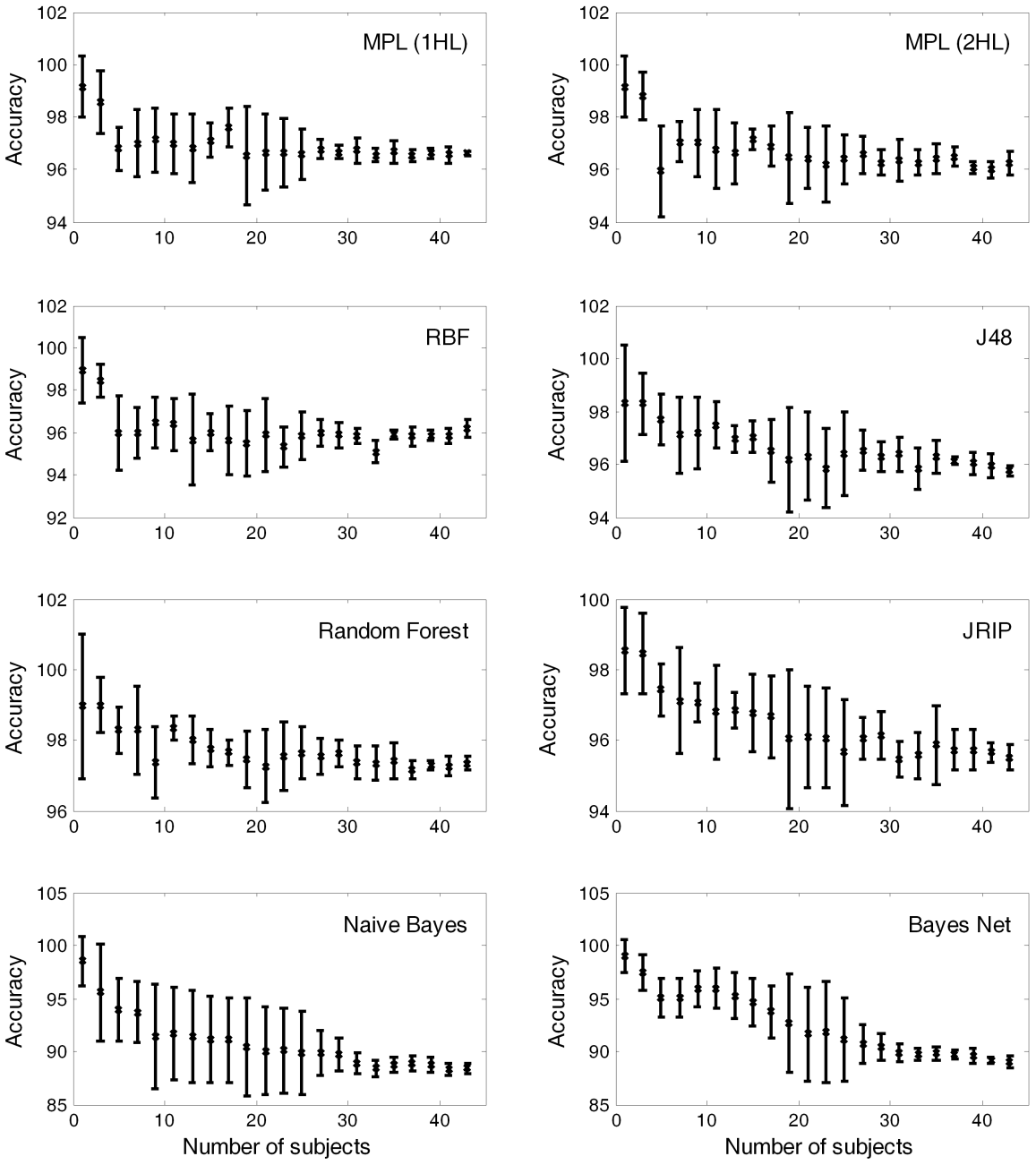

Figure 5.2: Classifier performance curves obtained by the variation of the number of subjects in the dataset.

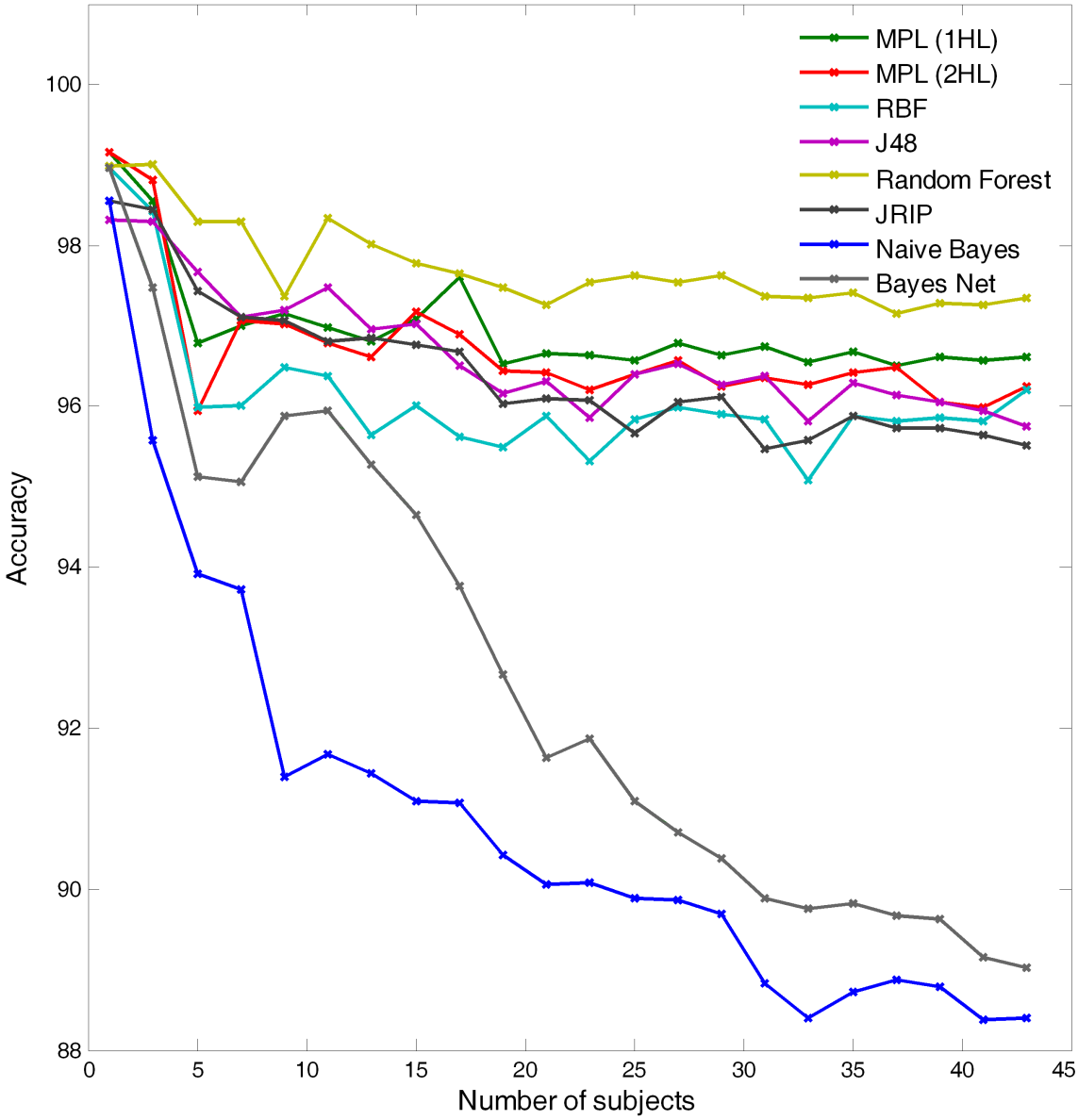

Figure 5.3: Accuracy measured with increased subjects number.

5.7.2.2 Artificial Neural Networks

Different parameter configurations were tested for each neural network (MPL-1 hidden layer, MPL- 2 hidden layers and RBF). The first trials showed that the manipulation of the neurons number is preponderant in the analysis. During the procedure, accuracy, sensitivity, specificity and precision were computed. The results are detailed below:

- **MPL-1 Hidden Layer**

  The analysis of Table 5.4 suggests that the optimal number of hidden neurons may be approximately between 19 and 20. The first condition (run 5) presents higher sensitivity and precision but, on the other hand, run 6 presents higher specificity. The second configuration is preferred for further analysis due to the higher specificity, which is quite important. One should consider models with higher specificity, because it is preferable having a classifier with few FPs as possible, to avoid the misdiagnosing of unhealthy patients as healthy. However, methods with high sensitivity should not be discarded, since high FNs rates can lead to the identification of other problems (such as, bad data acquisitions, for instance).

Table 5.4: MLP performance study for 1 hidden layer.

Run	Hidden neurons	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)
1	15	96.06	95.43	96.64	96.38
2	16	96.57	95.56	96.58	96.35
3	17	96.67	96.56	96.78	96.56
4	18	96.84	96.63	97.04	96.85
5	19	96.98	96.84	97.10	96.91
6	20	96.98	96.77	97.17	96.83
7	21	96.70	96.07	97.10	96.40
8	22	96.77	96.98	96.58	96.37
9	23	96.34	95.93	96.71	96.47
10	24	96.60	95.35	96.85	96.62
11	25	96.34	95.35	96.32	96.08

- **MPL-2 Hidden Layers**

  A similar approach was performed for the 2-hidden layer case, following the methodology used for MPL-1 hidden layers. The number of hidden neurons was manipulated in the $20-22$ range. The overall performance is listed on Table 5.5, where run 3 excels for the case of 22 considered neurons on the 2-layers.

Table 5.5: MLP performance study for 2 hidden layer.

Run	**Hidden neurons (layer1,layer2)**	**Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**	**Precision (%)**
1	20,20	96.51	96.56	96.45	96.22
2	21,21	96.44	96.42	96.45	96.22
3	22,22	96.74	96.52	96.91	96.70

- **RBF**

  The analysis of the RBF performance focused on two parameters: the clustering seed (a random estimate of the initial weight) and the number of clusters. Results are expressed in Table 5.6. It was verified that the number of clusters was preponderant. Run 6 presents the best global results for accuracy, specificity and precision. Therefore, in the runs 8 and 9, the same number of clusters was used and the clustering seed was manipulated. The seed value is used for generating a random number which is used for making the initial assignment of instances to clusters. Run 8 presents the best results, namely higher accuracy (96.23 %), sensitivity (96.23 %) and precision (96.65 %).

Table 5.6: RBF performance study, while manipulating clustering seed and the number of clusters.

Run	**Clustering seed**	**Number of Clusters**	**Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**	**Precision (%)**
1	1	2	91.92	92.56	91.33	91.92
2	1	5	92.26	90.46	93.96	92.26
3	1	10	94.49	94.25	94.71	94.49
4	1	20	95.32	95.02	95.60	95.32
5	1	50	95.79	94.46	97.04	95.79
6	1	100	96.03	94.74	97.24	96.03
7	1	250	95.32	94.53	96.06	95.32
8	5	100	96.23	96.23	95.79	96.65
9	10	100	95.76	95.12	96.36	96.13

5.7.3 Classifiers Selection

The classification results are presented in Table 5.7. These results demonstrate the superior performance of Random Forest. Classifiers with $90-95$ % accuracy are regarded as solid rules [15], which in this case represent six of the eight evaluated classifiers, and the others remain very close to 90 % (Table 5.7). On the other hand, ANN methods require high computational resources, as they can take several minutes to be trained. ANNs showed excellent results in the specificity analysis, which means that FP error rate is low.

The ROC curves corresponding to four of all the classifiers are plotted in Figure 5.4. The area values obtained for each of the classifiers were: 0.994 for Random Forest, 0.961 for J48, 0.965 for JRIP and 0.939 for BayesNet. Examining the ROC curves it is possible to conclude that the better results were achieved with Random Forest, J48 and JRIP.

Table 5.7: Comparative classifier selection performance results.

Classifier	**Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**	**Precision (%)**	**Training Time**
C4.5	95.72	95.58	95.86	95.58	**
Random Forest	97.15	97.47	96.85	96.66	**
RIPPER	95.72	95.72	95.34	95.07	***
Naïve Bayes	88.05	88.05	86.60	89.42	*
Bayesian Network	89.01	84.60	92.90	91.80	*
MLP (1 HL)	96.98	96.84	97.11	96.91	****
MLP (2 HL)	96.74	96.56	96.91	96.70	*****
RBF	96.23	96.23	95.23	96.93	***

* $[0-2\ s]$, ** $[2-10\ s]$, *** $[10-20\ s]$, **** $[20-60\ s]$, ***** $[>60\ s]$

5.7.4 Predictive Model Evaluation

The data prediction results (for GroupD) are presented in Table 5.8. For each subject, the class was predicted using the J48, Random Forest, BayesNet and JRIP classifier models previously tested. All the pulses were considered, and the final prediction is presented as a percentage of the number of class A and class B pulses, within the dataset.

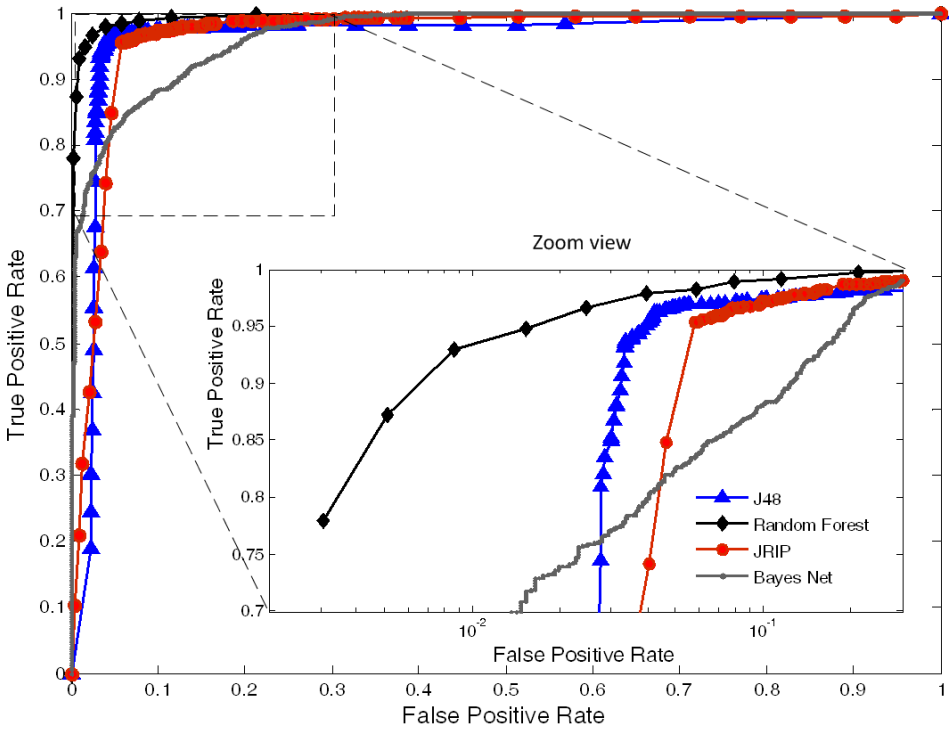

Figure 5.4: Semi-log ROC curves for J48, Random Forest, JRIP amd BayesNet classifiers.

The AIx values were compared with these labels. The correlation between the higher class A percentage and the higher positive AIx values is clearly visible (Figure 5.5). For the pulses in class A, an average AIx value of 14.80 ± 8.60 % was obtained and, for the class B group, a value of 4.10 ± 8.70 % was observed.

Table 5.8: Data obtained from class prediction analysis and the AIx values (for each subject).

Subjects	**Predicted class (%)**		**AIx (%)**
	A	**B**	
1	87.13	12.87	19.05 ± 3.57
2	86.09	13.91	12.28 ± 10.11
3	76.00	24.00	16.15 ± 3.76
4	65.53	34.47	7.13 ± 9.41
5	62.98	37.02	10.70 ± 11.58
6	61.85	38.15	12.46 ± 7.14
7	61.04	38.96	11.50 ± 8.40
8	60.66	39.34	9.89 ± 4.37
9	54.87	45.13	6.63 ± 9.92
10	13.95	86.05	-3.33 ± 8.29

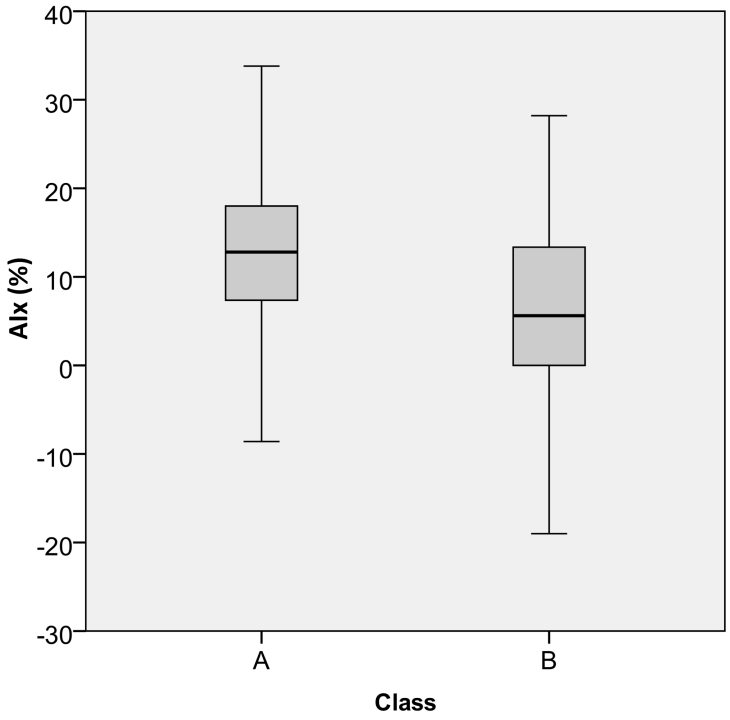

Figure 5.5: Boxplot of AIx distribution for classes A and B. The horizontal bold lines represent the medians, the box represent the interquartile ranges (50 % of the distribution) and the whiskers represent the range of values obtained for all subjects from GroupD in an overall set of 627 pulses.

5.7.5 Comparison of Risk Scores Estimates

The comparison of the risk values obtained for the GroupH by each of the scores was performed through on Bland-Altman plots (Figure 5.6) and Pearson's correlation matrix (Table 5.9).

- **SCORE *vs.* FRS**

  The mean risk of CHD events calculated by the SCORE function was 2.5 % per year and by the FRS function was 12 % per year. A significant correlation between these methods $(r = 0.641,\ p < 0.01)$ was verified. The Bland-Altman plot (Figure 5.6(a)) showed a systematic mean difference of -9.50 % that resulted from the overestimation of risk levels by the FRS.

- **SCORE *vs.* PROCAM**

  Estimates of CHD event risk by the SCORE function (2.5 % per year) were much lower than those obtained by PROCAM (17.31 % per year). It was observed a moderate correlation

between these methods ($r = 0.639,\ p < 0.01$). The Bland-Altman plot also shows a large, but systematic difference (mean value of -14.81 %), such as it is visible in Figure 5.6(b).

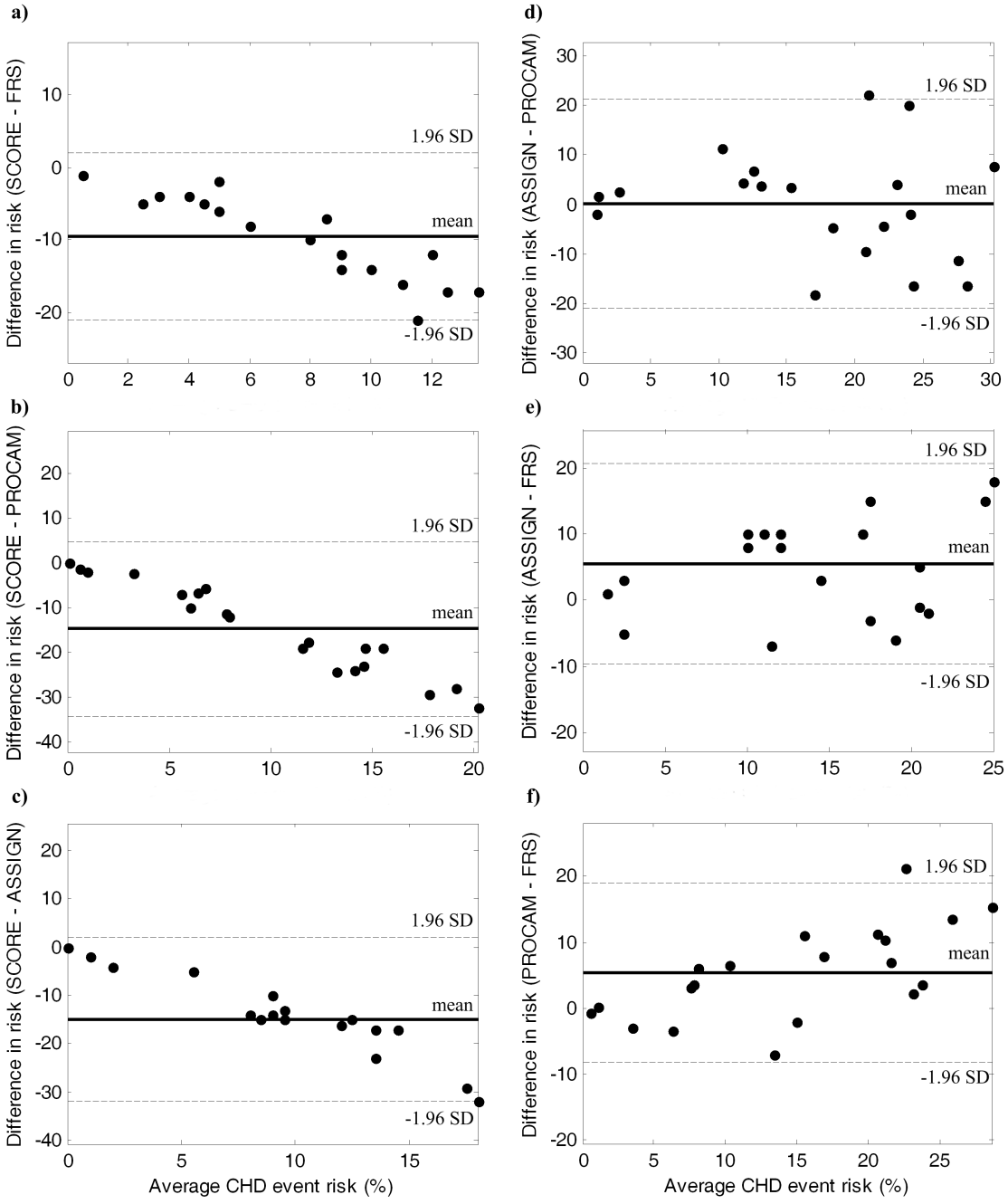

Figure 5.6: Bland-Altman plots for the individual risk estimative by the FRS, PROCAM, ASSIGN and SCORE functions.

- **SCORE *vs.* ASSIGN**

  Estimates of CHD event risk by the SCORE function were also much lower than those by ASSIGN (17.50 % per year). In this case a weak but significant correlation between estimates was achieved ($r = 0.529,\ p < 0.05$). The Bland-Altman plot (Figure 5.6(c)) confirmed the large and systematic difference, among methods on average (-15.00 %).

- **ASSIGN *vs.* PROCAM**

  The mean risk of CHD events calculated by the ASSIGN function was 17.50 % per year, and by the PROCAM function was 17.31 % per year, presenting only a mean difference of 0.19 %, the lowest difference verified within all the considered scores. A moderate correlation for the CHD risk estimates from both methods was verified ($r = 0.588,\ p < 0.01$). The Bland-Altman plot (Figure 5.6(d)) showed agreement on average, and does not evidence a systematic bias, but presents large variability at higher values.

- **ASSIGN *vs.* FRS**

  The mean risk of CHD events calculated by the ASSIGN function was 17.50 % per year, and by the FRS function was 12 % per year, with a mean difference of 5.50 %. A moderate correlation between CHD risk estimates among methods ($r = 0.670,\ p < 0.01$) was verified. The Bland-Altman plot (Figure 5.6(e)) does not show any systematic error.

- **PROCAM *vs.* FRS**

  The mean risk of CHD events calculated by the PROCAM function was 17.31 % per year, and by the FRS function was 12 % per year. The higher significant value was verified for these methods ($r = 0.782,\ p < 0.01$). However, from the analysis of Figure 5.6(f) it was also verified that the PROCAM function comparatively to the FRS tends to overestimate the CHD risk in general (mean difference of 5.31 %), in spite of underestimation at lower risk levels. Similar results were achieved, by Haq *et al.* [202].

Table 5.9: Pearson's correlation matrix obtained between all variables and risk scores.

	DBP (mmHg)	**Total-CH** (mg/dL)	**CH-HDL** (mg/dL)	**TGL** (mg/dL)	**RWTT** (ms)	**RWA** (a.u.[a])	**SWTT** (ms)	**LVET** (ms)	**AIx** (%)	**FRS** (%)	**SCORE** (%)	**ASSIGN** (%)	**PROCAM** (%)
SBP	0.812**	0.357	0.260	0.194	0.335	0.149	0.363	0.358	-0.158	0.177	0.217	0.089	0.284
DBP	-	0.408	0.211	0.374	0.293	0.166	0.183	0.214	-0.178	0.167	0.202	0.051	0.465*
Total-CH	-	-	0.133	0.718**	0.081	-0.005	0.330	0.374	-0.054	0.313	-0.014	-0.047	0.250
CH-HDL	-	-	-	0.214	0.017	0.283	-0.074	-0.002	-0.321	-0.456*	-0.380	-0.581**	-0.267
TGL				-	-0.140	-0.070	0.057	0.203	-0.004	0.060	0.002	-0.237	0.159
RWTT					-	0.708**	-0.283	-0.158	-0.758**	-0.016	0.076	-0.073	0.102
RWA						-	-0.589**	-0.486*	-0.965**	-0.353	-0.151	-0.511*	-0.271
SWTT							-	0.928**	0.642**	0.580**	0.251	0.346	0.323
LVET								-	0.493*	0.606**	0.330	0.308	0.388
AIx									-	0.384	0.161	0.491*	0.257
FRS										-	0.641**	0.670**	0.782**
SCORE											-	0.529*	0.639**
ASSIGN												-	0.588**
PROCAM													-

Significant levels: $**p < 0.01$, $*p < 0.05$

[a] Arbitrary Amplitude Units

5.7.6 PWA Parameters Evaluation

The analysis of Table 5.9 allows to conclude that significant correlations occur between the AIx and the remaining parameters, namely strong correlations with RWA ($r = -0.965,\ p < 0.001$) and RWTT ($r = -0.758,\ p < 0.001$). The negative values in the correlation are associated to early reflections that tend to occur at lower amplitudes. The AIx values present a significant correlation with SWTT $(r = 0.642,\ p < 0.001)$ and LVET $(r = 0.493,\ p < 0.001)$. On the other hand, no significant correlations were verified between RWTT and SWTT ($r = -0.283,\ p = 0.227$) and between RWTT and LVET ($r = -0.158,\ p = 0.506$). For an early reflection wave, the SWTT and LVET values increase, leading to the negative values observed. Finally, a strong significant correlation was also achieved between SWTT and LVET ($r = 0.928,\ p < 0.001$).

Figure 5.7 shows the correlation plots for SWTT, RWTT and LVET parameters, with each risk scores (FRS, SCORE, ASSIGN and PROCAM). A clearly positive trend occurs for the SWTT and LVET. This result supports the assumption that the forward wave duration is associated to the CHD risk increase, as was also demonstrated by Mitchell el. al. [140]. RWTT values assume a low relevance in the risk prediction. This conclusion is also supported by the low significant values observed in Table 5.9.

It was also observed an interesting correlation value between the AIx values and the score results, namely for the ASSIGN ($r = 0.491,\ p < 0.05$), and FRS ($r = 0.384,\ p = 0.095$). Similar results were achieved by Song *et al.* [204].

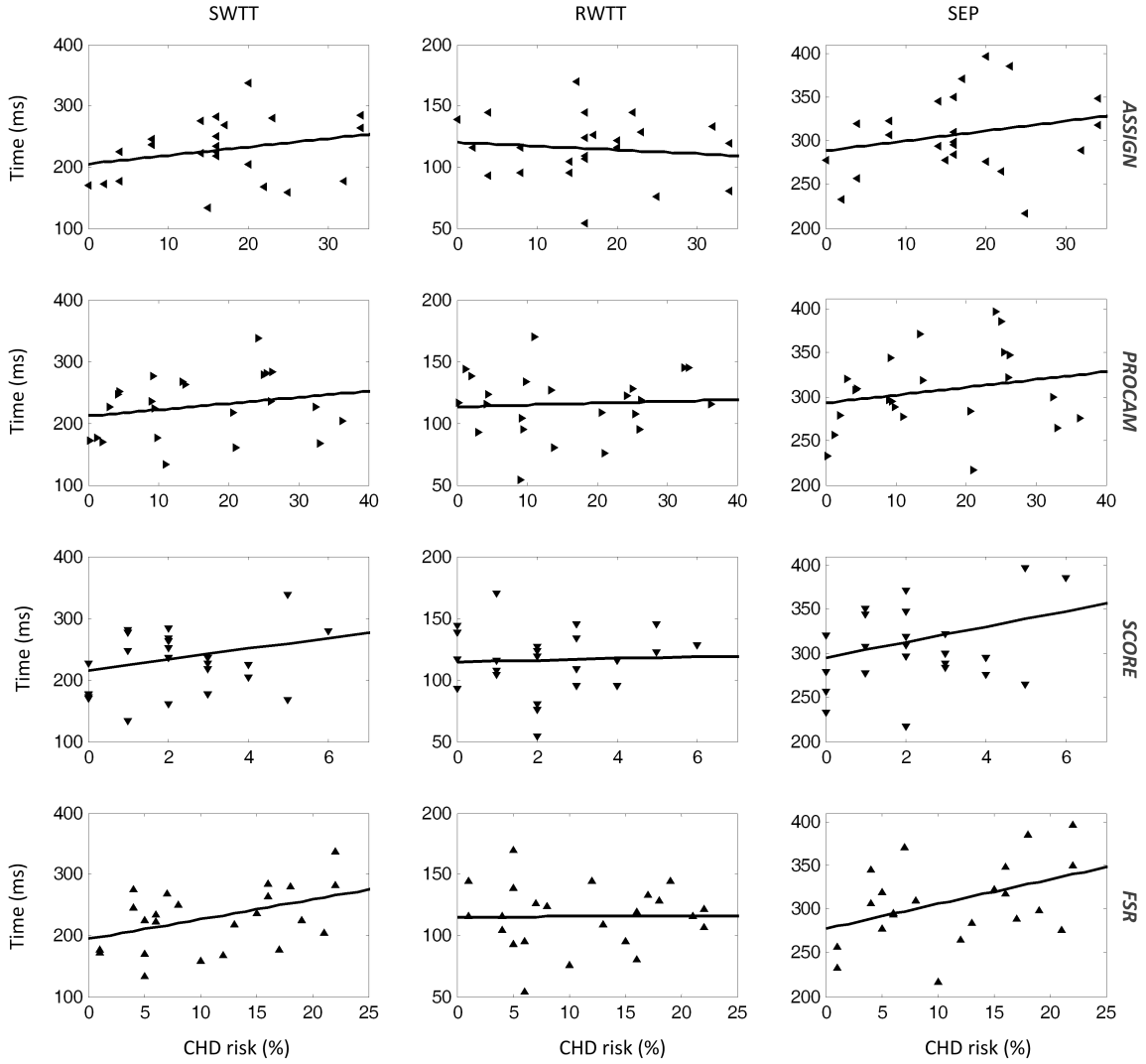

Figure 5.7: Correlation plots between transit time based parameters and all the considered risk scores (ASSIGN, PROCAM, SCORE and FSR).

5.7.7 Clustering Analysis

The cluster analysis was performed for GroupC and GroupH. As referred previously, different nomenclatures were adopted to help understanding the data at hand, numeric labels ($1,\ 2,\ ...$) for the GroupC and alphabetic labels ($A,\ B,\ ...$) for the GroupH.

5.7.7.1 GroupC

The first approach consisted in the selection of the best clustering algorithm to deal with the set of features at test, as well as the ideal number of clusters to the group characterization. Figure 5.8 displays RWTT and SWTT plot for 2-clusters analysis using EM (a) and *k*-means (b) algorithms. Red and blue coloured points represent the pulse labels (blue = Cluster 1, red = Cluster 2).

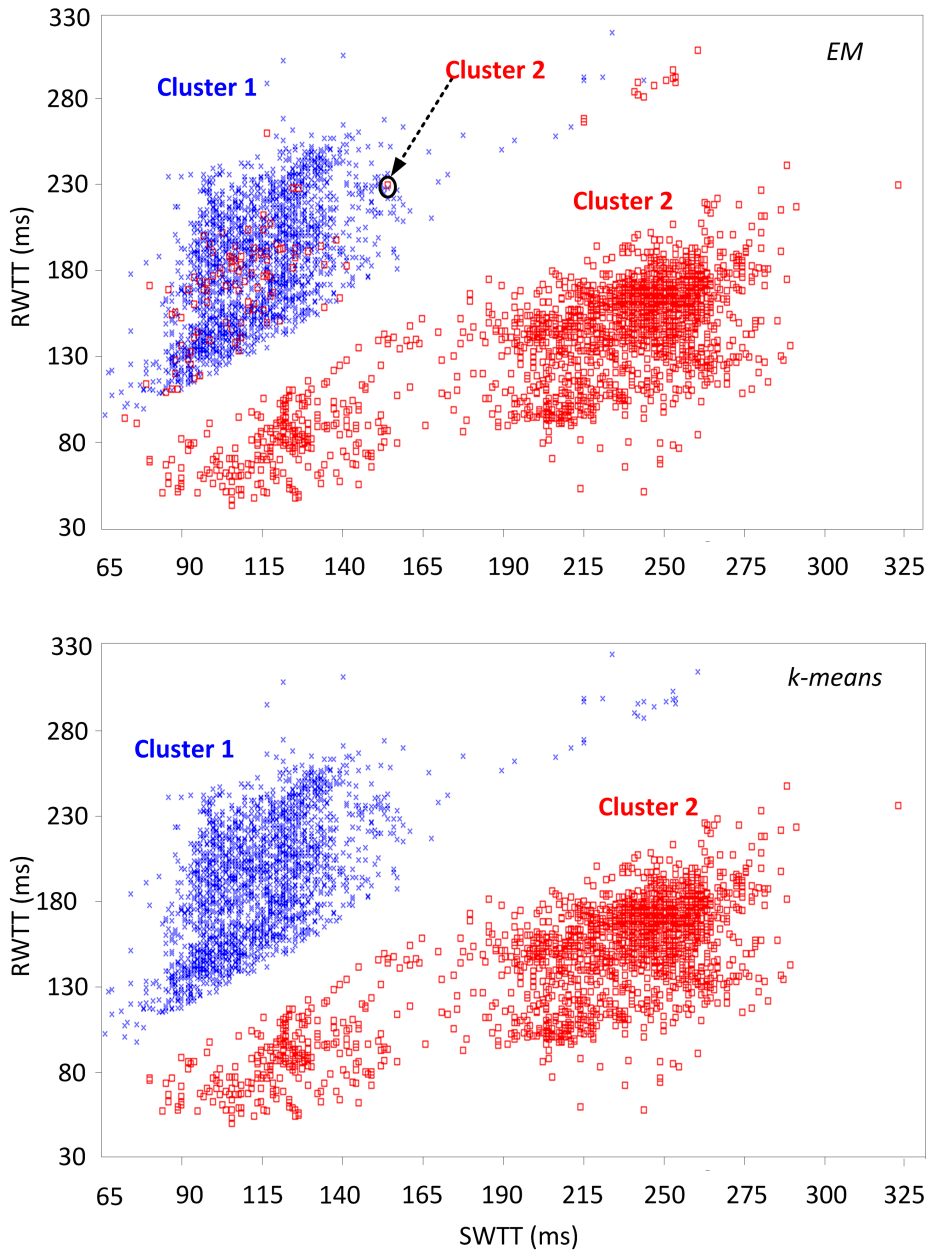

Figure 5.8: Clusters performance obtained for EM (top) and *k*-means (bottom) algorithms using two clusters, Cluster 1=blue, Cluster 2=red.

Categorical features (gender, smoker) were not considered during the analysis, since *k*-means analysis is not able to deal with this kind attributes.

The figure shows that the EM plot has an unsatisfactory division between the clusters, with some points from Cluster 2 (red) identified as being in the Cluster 1 area (blue), as well. Visually, the results from the *k*-means clustering have a more efficient separation, as the dataset is partitioned in two homogeneous risk groups. For both EM and *k*-means, the blue group (Cluster 1) represent the pool of healthier subjects when compared to the red group (Cluster 2), since it represents the cases where the reflection wave arrives after to the systolic peak. The analysis of Cluster 2 also indicates the presence of another sub-cluster. Taking this into account, the *k*-means method was used to explore the distribution of a third cluster, since it performed well, comparatively to the EM, in the two cluster distribution. The obtained results for a 3 cluster distribution are presented in Figure 5.9. The detailed information about this cluster disribution is presented in Table 5.10. The Cluster 3 (green homogeneous zone) is mostly representative of Type C APW pulses (see waveform types definition in Section 4.4), where $RWTT > SWTT$. The mean AIx value of the cluster centroid is negative (-11.2 %), thus being considered the lower risk cluster. Cluster 2 (red points) predominantly consists of Type B pulses, with $SWTT > RWTT$ and $AIx > 0$, being thus considered the intermediate risk group. Cluster 1 pulses (blue points) represent the less homogeneous group,

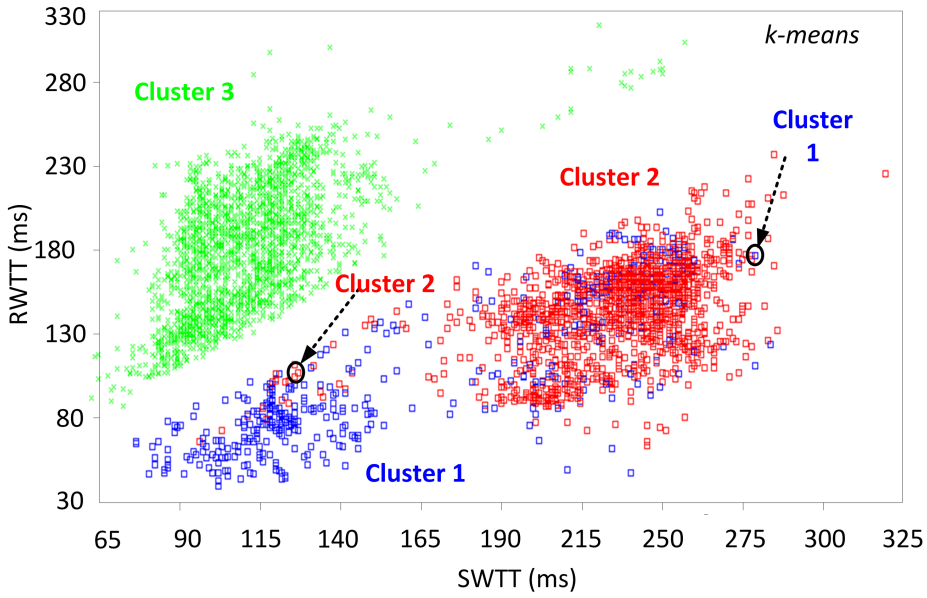

Figure 5.9: RWTT and SWTT plot after *k*-means clustering for using three clusters. Blue=Cluster 1, red=Cluster 2 and Green=Cluster 3.

Table 5.10: Average values for the 3-clusters features.

Attributes	Cluster		
	1	2	3
Pulses	458	1550	2463
Age (y)	21.6	21.0	21.7
Weight (kg)	63.0	55.4	63.2
Height (m)	1.7	1.6	1.7
BMI (kg/m^2)	21.3	20.8	21.8
SBP (mmHg)	109.9	106.1	108.6
DBP (mmHg)	69.6	70.3	68.9
HR (bpm)	72.8	67.5	72.9
SWTT (ms)	172.7	234.4	117.1
RWTT (ms)	103.0	143.2	179.9
LVET (ms)	240.3	306.4	274.4
SWA (a.u. [a])	1.0	1.0	1.0
RWA (a.u. [a])	0.8	0.9	0.9
DWA (a.u. [a])	0.8	0.8	0.7
R1 (-)	67.6	72.0	157.3
R2 (-)	0.8	0.8	0.7
R3 (-)	0.2	0.1	-0.1
R4 (-)	0.8	0.9	-0.9
AIx (%)	21.0	12.6	-11.2
$RMSSD_{SPt}$ (ms)	34.9	25.5	21.7
$RMSSD_{SPA}$ (a.u. [a])	0.0	0.0	0.0
$RMSSD_{Pia}$ (a.u. [a])	0.5	0.0	0.2
$RMSSD_{DWT}$ (ms)	54.2	18.5	42.8
$RMSSD_{DWA}$ (a.u. [a])	0.1	0.1	0.1
FWHM (ms)	459.0	450.6	466.3
RMSE (%)	0.1	0.0	0.1

[a] Arbitrary Amplitude Units

with some points also scattered across the Cluster 2 area. These pulses are mainly APW Type A pulses, with some punctual Type B pulses. This group evidences a higher cardiovascular risk comparatively to the Cluster 2 and Cluster 3.

During the clustering procedures, $RMSSD_{Pit}$ negatively interfered with cluster measurements. For this reason, this parameter was removed from the analysis alongside the categorical features.

5.7.7.2 GroupH

The *k*-means was also used in the GroupH analysis. Figure 5.10 depicts the 2-cluster distribution, where Cluster A and Cluster B were the labels adopted.

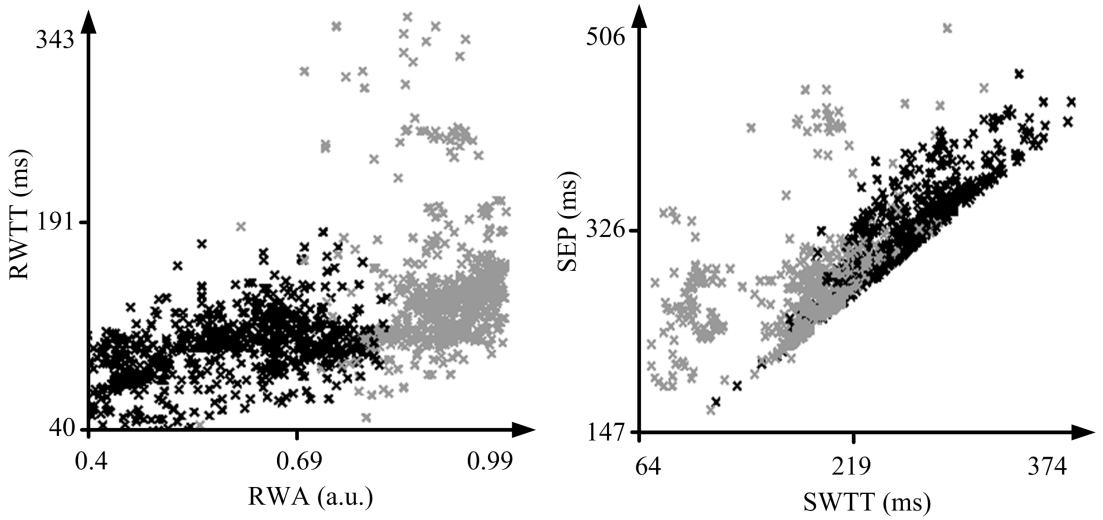

Figure 5.10: Scatter plots for: (a) RWTT and RWA, (b) LVET and SWTT, where the grey and black markers denote Cluster A and Cluster B, respectively.

The characteristics of each of the clusters are presented in Table 5.11. It can be verified that the clusters do not differ significantly according to the SBP and DBP values. The most significant

Table 5.11: Cluster distributions for subjects in GroupH.

Attributes	**Cluster A**	**Cluster B**
SBP (mmHg)	169.68±10.38	172.56±13.59
DBP (mmHg)	101.22±7.03	102.00±7.14
Total-CH (mg/dL)	206.50±26.16	204.01±13.69
CH-HDL (mg/dL)	59.35±25.07	66.33±23.31
TGL(mg/dL)	160.41±44.95	142.94±40.96
RWTT (ms)	99.90±23.62	140.32±42.76
RWA (a.u. [a])	0.60±0.11	0.90±0.06
SWTT (ms)	257.34±35.76	185.34±40.70
LVET (ms)	326.24±39.43	272.16±44.17
AIx (%)	40.01±11.15	6.2±9.67
Clustered instances	723 (52 %)	664 (48 %)

[a] Arbitrary Amplitude Units

parameters (except APW parameters) were the CH-HDL and the TGL values. Cluster A presents higher TGL values, but also lower CH-HDL levels, that are characteristics of subjects at risk. From the waveform analysis, lower RWTT, higher SWTT and, consequently, higher positive AIx values were obtained. The suitable number of clusters was considered to be two as may be confirmed by visual inspection.

During the evaluation, the pulse instances were independently grouped according to their cluster similarities, being the maximum level of similarity achieved, for each subject, when all the pulses fit in a single cluster. So, for each subject a value representing the cluster risk (as percentage) was computed and used in the correlation with other available parameters (including the studied scores). Since only two clusters are studied, at this point, the values of Cluster A (%) are symmetric to the values of Cluster B (%). The Cluster A (%) correlations are presented in Table 5.12.

The lipidic and the BP values present low significance with the Cluster A (%), similarly to the values observed from the correlation with scores (see Table 5.9). On the other hand, significant correlation values were obtained between Cluster A (%) and PWA parameters. Additionally, it

Table 5.12: Spearman's correlation coefficients obtained for Cluster A (%).

Attributes	**Cluster A (%)**
SBP	-0.131
DBP	-0.080
Total-CH	0.011
CH-HDL	-0.177
TGL	0.017
AIx	0.921 **
SWTT	0.649 **
RWTT	-0.632 **
RWA	-0.933 **
LVET	0.574 **
FRS	0.458 *
SCORE	0.275
ASSIGN	0.582 **
PROCAM	0.391

Significant levels: ** $p < 0.01$, *$p < 0.05$

was obtained a significant correlation between Cluster A (%) and the risk scores. In this case was observed a very significant correlation with ASSIGN ($r = 0.582,\ p < 0.01$), a significant correlation with FRS ($r = 0.458,\ p < 0.05$) and, weaker correlations with SCORE ($r = 0.275,\ p = 0.241$) and PROCAM ($r = 0.391,\ p = 0.088$). These results are good indicators for the use of this methodology as a tool for the cardiovascular risk assessment.

5.7.7.3 Comparison of Groups

The evaluation was performed by the comparison of the clusters in GroupH (Cluster A and Cluster B) and clusters in GroupC (Cluster 1, Cluster 2 and Cluster 3). The cardiovascular risk, assessed by the scores is available for the GroupH, for that it was demonstrated that a higher risk is association to the Cluster A. Since, there is no information about the risk associated to GroupC, each of the three clusters was compared to the GroupH parametrizations. The comparison was performed using the chi-squared (χ^2) value, obtained from the Kruskal-Wallis test, as shown in Table 5.13.

There are significant differences for the majority of clusters in GroupH and GroupC, as expected due to the distinct population characteristics of each group. However, some similarities were found between the Cluster B (the score at lower risk in groupH) and the Cluster 1 (belonging to the GroupC), namely: for LVET ($\chi^2 = 4.79,\ p < 0.01$) and SWTT ($\chi^2 = 0.02,\ p < 0.01$) measures. Additionally, from the analysis of Table 5.10 that Cluster 1 has a higher AIx value associated, and

Table 5.13: Comparison of clusters from GroupH (Cluster A and Cluster B) and GroupC (Cluster 1, Cluster 2 and Cluster 3) using Kruskal-Wallis test, measured through χ^2 value.

GroupH	**GroupC**	**Parameters**					
		RWTT	*RWA*	*SWTT*	*RWTT*	*LVET*	*AIx*
ClusterA	Cluster1	38.66 **	402.42**	526.65**	38.66**	369.23**	402.40 **
	Cluster2	1149.82**	1399.96 **	248.87 **	1149.82**	154.81**	1399.97**
	Cluster3	1454.05**	1545.50**	1619.16**	1454.05**	706.65**	1638.37**
ClusterB	Cluster1	536.38 **	458.25**	0.02	536.38**	4.79	538.34 **
	Cluster2	265.77**	3.16 **	892.34 **	265.77**	598.03**	49.97 **
	Cluster3	593.88**	4.06**	998.62**	593.88**	60.22**	1158.33**

Significant levels: ** $p < 0.01$

that it is the cluster at higher risk in GroupC. So, it can be concluded, that this cluster presents closest parametrizations with the ClusterB from hypertensive group, a cluster at risk, but in lower levels. It would not be trustable, if the similarities occur for the group of subjects in advanced stages of disease. This conclusion is very important, since allowed the screening of a healthy population (using only waveform parameters) concerning the cardiovascular risk using clustering methodologies.

5.8 Discussion

This section includes the discussion of the results obtained with the data mining implemented methodology for the feature selection, classification, clustering and predictive model algorithms, as well as risk scores implementation.

5.8.1 Features Selection

Concerning the predictive value of the APW attributes (Figure 5.1), arrival time (RWTT) revealed to be the most significant in the ”subjects” discrimination. This relevance of Pi-related parameters can be explained by the known fact that arterial stiffness strongly correlates to its components (both in time and amplitude). Clinically, this effect is usually quantified on the well-known AIx. Compared to the single AIx analysis, the proposed framework has the main advantage of being more flexible when dealing with poor quality pulses, where pulse wave analysis can be difficult, or even impossible, to perform. The difficulty in obtaining AIx values invariably occurs in cases where the inflection point is very close to (or buried in) the SP. In these cases, this approach is an efficient way to sort out this ambiguity by the combination of the information coming out from several other morphological parameters. Naturally, for the developed model, some degree of dependence on the accuracy of the inflection determination remains. However, the interaction with other parameters, implied by the method, decreases this dependence to a point where prediction becomes reliably possible. Naturally, systolic time (SWTT) also appears as one of the key attributes in this dataset. The other indexes, R1, R2 and R3, still have considerable merit, but not as high as those of Pi and SP. However, they are quite important, because being Pi-independent parameters, they account for

a more robust class prediction, whenever Pi is hidden in the APW and its parametrization is not achieved. Dicrotic wave values changed less than expected, due to the absence of registered cardiac valves complications in this database. Also related to this fact is observed the low predictive value of the R4 ratio.

5.8.2 Classification Model

The results from the classification performance are expressed in Table 5.7. All classification methods except Bayesian based classification exhibited accuracy values >95 %. Random Forest was the best classifier, in overall, since it produced the highest accuracy and sensitivity values. In comparison with the other classifiers, ANN methods require higher computational resources, as they can take more than 30 seconds to be trained. However, ANN algorithms have shown superior results in terms of specificity, which means that the FP error rate is low. The three classifiers with highest accuracy scores were: Random Forest, MLP-1 hidden layer and MLP-2 hidden layers.

Results from the ROC curves confirm the good accuracy verified previously. A very high true positive rate was achieved for Random Forest, J48 and JRIP contrasting with the BayesNet performance. These differences are clear in the zoom window view, in Figure 5.4, where the x axis is plotted on a logarithmic scale.

The analysis of the learning curves revealed that it is possible achieve the classifiers maximum performance with $25-30$ subjects in the dataset. It was also concluded that the combination of several classifiers should be used in APW studies in order to avoid the limitations of single-parameter analysis.

5.8.3 Predictive Model

In the validation group, the subjects belonging to GroupD are young (average of 24 years old) but also present positive AIx values, out of the normal range for this age [144], varying from -3.33 % to 19.05 %. The application of the previous model intends to discriminate the patterns and the relationships in the pulse morphology and the identification of interesting risk labels. These values were validated by AIx values. A positive correlation between risk (class A in this dataset) and AIx

positive values was verified. The knowledge obtained through this study provides more information than typical information provided by pulse wave analysis techniques.

5.8.4 Risk Scores

The first point concerning attention is related with the application of different risk scores in the same population. From the analysis of all risk scores, it is visible the underestimation of risk levels by the SCORE, comparatively to the other tools. Higher errors tend to occur for patients at higher risk, but the errors seem to be systematic. The values from the correlation analysis are in agreement with other authors, such as the values obtained by Haq *et al.* [202], that observed a significant strong correlation ($r = 0.82,\ p < 0.001$) between FRS and PROCAM, similar to the value obtained in this study ($r = 0.782,\ p < 0.001$). A significant correlation was also verified between ASSIGN and FRS ($r = 0.670,\ p = 0.001$), but lower than the values obtained by Woodward *et al.* [205], that for a population under gender discrimination obtained a correlation of 0.92 for men and 0.90 for women. The verified differences within all scores are acceptable and justified by the dissimilarities of the populations (age, the duration of the follow-up, number or participants and the definition of end-point). Risk prediction, inevitably, has associated errors, but the use of multiple scores could contribute to enhance its accuracy.

Parameters used by each risk scores are similar: age is included in all scores, lipidic measurements have been used by several authors, being the Total-CH and the CH-HDL the most used. Inclusion of other parameters, such as family history or smoking habits can improve the results, and are effectively used by most of the scores. Additionally, diabetes information is required by the ASSIGN and PROCAM scores, and can be particularly valuable in young people. Generally, all of these parameters are positively correlated to the risk scores with the exception of CH-HDL that correlates negatively to the scores. The CH-HDL values are strongly correlated with FRS ($r = -0.456,\ p < 0.05$) and ASSIGN ($r = -0.581,\ p < 0.01$).

Significant correlations were verified for the PWA parameters. The SWTT and LVET parameters presented higher values comparing to the reflected wave components (RWTT and RWA). The lower significance verified for the reflected wave components can be interpreted as part of the more complex reflection process. The reflection wave results from the sum of waves at the innumerous

reflection sites and not from a single effective site. Analysing each one of the parameters, the correlation between RWA and SWTT ($r = -0.589,\ p < 0.01$) and the correlation between RWA and LVET ($r = -0.486,\ p < 0.05$) were preponderant comparatively to the values observed between RWTT and SWTT ($r = -0.283,\ p > 0.05$) and between RWTT and LVET ($r = -0.158,\ p > 0.05$). Additionally, strong correlations for the AIx and RWA ($r = -0.965,\ p < 0.001$) and for the AIx and RWTT ($r = -0.758,\ p < 0.001$) were verified, as expected.

5.8.5 Clustering Analysis

The information that is extracted from the clustering analysis can be crucial to fully understand of the data, mainly when there is no, or little, available information. It was verified, that Total-CH and TGL values are intrinsically related to the APW variables, and than an increase of these levels is associated to the RWTT decrease and SWTT increase. On the other hand, the higher CH-HDL values are associated to the increase of RWTT and decrease of SWTT values. A significant correlation for the cluster output with the ASSIGN ($r = 0.582,\ p < 0.01$) and with FRS ($r = 0.458,\ p < 0.05$) was also verified.

This method is particularly interesting, if it was considered that this approach can avoid the requirements of the classification procedures which often require costly labelling of a large set of patterns, such as biochemical analysis used by traditional risk scores, as demonstrated on the cluster analysis of subjects in GroupC. It was possible to identify the similarities with GroupH clusters, using only morphological parameters. The possibility to emulate the performance of the traditional risk scores by the use of only non-invasive parameters, which can be easily obtained using a PWA technique, is an efficient alternative approach to study large datasets. In this sense, clustering analysis is an easy method for large populations screening producing valuable knowledge for posterior prioritising people for pharmacological measures. In summary, from the analysis of the given results, it can be concluded that clustering risk assessment performed with multiple PWA parameters is correlated to the risk score values.

5.9 Conclusion

The healthcare public and private system faces the challenge of an increasingly ageing population and the escalation of medical costs. Therefore, the development of techniques that allow for earlier identification of subjects at risk, by including customised predictive and preventive methodologies, are essential.

The developed methodology has proven its usage in screening studies due to its descriptive and predictive power. The successful application of data mining techniques can help to predict under-diagnosed patients, and identify and classify people at-risk in terms of cardiovascular health, as well as reduce healthcare costs. In clustering methods, more than the usefulness when there is no, or little, information available, the information that is extracted from them can be crucial in fully understanding the data.

The high limitation of this work is its cross-sectional design that can explain why PWA measures were low significantly associated with all cardiovascular risk outcomes. A prospective design would more effectively elucidate the relationship between arterial stiffness indices and the outcomes from the traditional scores. The sample size of the studied populations should also increase and include the stratification by medication use, age, diabetes duration, and/or gender.

6

Validation Tests

This chapter intends to assess the potential clinical use of the APW module. The validation tests consisted in: comparison of the parameter values for different groups; comparison of non-invasive APW measurements with intra-arterial measurements in patients, before and after an angioplasty procedure; and, assessment of the system repeatability (intra-operator variation), reproducibility (inter-operator variation) and pulse-by-pulse variability.

6.1 Motivation

The motivation of this work consisted in the validation of the APW module measurements. Generally, the validation of a new measurement device comprises the evaluation of its accuracy and precision, with measurements close to, and tightly clustered around the known value. The accuracy and precision of a measurement process is usually established by repeatedly measuring several parameters. The accuracy of a measurement system is the degree of agreement of measurements with a standard measurement methodology. The precision of a measurement system, also called reproducibility or repeatability, is the degree to which repeated measurements, under unchanged conditions, show the same results. In this sense, repeatability is the term used to describe the intrinsic precision of an instrument. The most reliable way to evaluate this parameter is through the repetitive analysis over a short period of time of the same measurand. This procedure implies the same conditions of measurement, namely, the same measurement procedure, the same observer, the same measuring instrument and, the same measurement location. On the other hand, reproducibility is the variability associated to different operators during the measurement of the same parameters.

6.2 Methods

The groups of subjects, described in Section 4.1, were used in the validation tests:

- The GroupC, GroupD and GroupH were used in the characterization of the distribution of values for the different parameters.
- The groupA was used for the comparison of the invasive and non-invasive waveforms.
- The groupR was used in the analysis of the significance of differences between consecutive sessions, different operators and subjects, and left and right carotid arteries. This group was also used in the analysis of the inter and intra-observer variability.

Statistical analysis was performed with SPSS v.17.0 (SPSS Inc, Chicago, IL, USA). The analysis comprised several tests, that are described in the next sub-sections. The results were considered statically different for *p-values* lower than 0.05. All data are expressed as $mean \pm SD$.

6.2.1 Descriptive Statistics Analysis

Descriptive statistics gives information about the number of people or cases in the sample, the number and percentage of males and females, the mean age of subjects, the smoking habits, as well as the range of values, thus providing a useful description of the dataset. The main measurements include mean and SD values, that are respectively, the effective methods for central trend and dispersion evaluation.

6.2.2 Normality Assessment

Before choosing a proper statistic test, it is important to check, if the variables in the dataset have a normal or a non-normal distribution. A normal distribution is used to describe a symmetric and bell-shaped curve, which has the greatest frequency of scores in the middle, while the smaller frequencies occur in the extremes [206]. The best quantitative way to assess the normality of the dataset is through the application of the Kolmogorov-Smirnov one-sample test [207].

6.2.3 Parametric and Non-Parametric Tests Application

There are several families of statistical techniques that can be used for the exploration of differences and relationships between groups. Generally, these techniques can be organized in parametric and non-parametric tests that should be properly chosen [206]. Parametric methods can produce more accurate and precise estimators and they have higher statistical power. However, before to the application of a parametric tests, one should have to consider some general assumptions, and if one or more are violated, the use of a non-parametric test must be considered. The general conditions are:

- **Normal distribution** - It is important the assessment of the data normality, since some parametric tests are not usually applicable for non-normal data. Fortunately, most of the parametric techniques are reasonably robust, and with large sample sizes, the violation of this assumption does not cause problems.

- **Continuous scale** - Parametric approaches assume that the dependent variables are measured in a continuous interval, and do not use discrete values.
- **Independence of observations** - Each observation or measurement must not be influenced by any other observation or measurement.
- **Random samples** - Parametric techniques assume that the scores are obtained using a random sample of the population. However, this is often not the case in real-life research.
- **Homogeneity of variance** - Parametric techniques also make the assumption that samples are obtained from populations with similar variability.

6.2.4 Exploration of Relationships Among Variables

The strength of the relationships between variables is an important topic that can deliver important information about the dataset. Correlation methods are used to describe the strength and direction of the linear relationship between variables [206]. The main tests used in the correlation assessment are the parametric test Pearson product-moment correlation, and the non parametric test Spearman rank order correlation.

A positive correlation indicates that when one variable increases, so does the other, whereas a negative correlation indicates that when one variable increases, the other decreases. Correlation coefficients can range from -1 to $+1$, with the size of the absolute value providing an indication of the strength of the relationship. A perfect correlation of -1 or $+1$ indicates that the value of one variable can be exactly determined by knowing the other. On the other hand, a correlation of 0 indicates the absence of relationship between the two variables [206]. It is important to understand, how to interpret the range of values, and to recognize small, medium and high correlations. The guidelines suggested by Cohen [208] are presented in Table 6.1.

Table 6.1: Correlation coefficient ranges according to Cohen *et al.* [208].

Correlation coefficient range			**Correlation**
$r = 0.10$ to 0.29	or	$r = -0.10$ to -0.29	Small
$r = 0.30$ to 0.49	or	$r = -0.30$ to -0.49	Medium
$r = 0.50$ to 1.00	or	$r = -0.50$ to -1.00	Large

6.2.4.1 Pearson Product-Moment Correlation

The Pearson's correlation test is a parametric test which was designed for linear detection of relationships among continuous variables. It can also be used in the presence of one continuous and one nominal variable [206].

6.2.4.2 Spearman's Rank Order Correlation

Spearman's rank order correlation is also used to calculate the strength and direction of the relationship between two variables, without making any assumptions about the frequency of the distribution of the variables. It is the non-parametric alternative to the Pearson correlation, and should be used in the case of highly skewed data or, when there are evidences of monotonic relationships [206, 209].

6.2.5 Comparison of Groups

Comparison of groups is another family of statistics tests, that can be used to find out statistical significant differences among groups. There are several tests that can be used depending of the number of variables. The parametric and non-parametric tests adopted are summarized in Table 6.2.

Table 6.2: Summary of the techniques used in the exploration of the relationships among variables.

Number of variables	Parametric	Non-Parametric
2	Independent-samples t-test	Mann-Whitney U Test
3	One-way ANOVA	Kruskal-Wallis Test

6.2.5.1 Independent Samples T-Test

The independent samples t-test is used to compare the mean scores of the two different groups of parametric data. The test reveals the statistical differences by the analysis of the mean score of the groups [206].

6.2.5.2 One-Way ANOVA

A one-way analysis of variance (ANOVA) is commonly used for the mean values comparison of more than two groups, in the presence of parametric data. It compares the variability between two different groups with the variability within the groups. An F ratio is calculated, which represents the variance between the groups. A large F ratio indicates that there is more variability between the groups (caused by the independent variable) than within each group (referred to as the error term) [206].

6.2.5.3 Mann-Whitney U Test

This technique is the non-parametric equivalent to the independent samples t-test, and it is used to evaluate the differences between two independent groups. This test compares the median values, instead of mean values [206].

6.2.5.4 Kruskal-Wallis Test

The Kruskal-Wallis H test is also a non-parametric alternative to the one-way ANOVA test, where the scores are converted to ranks, and the mean rank of each group is compared [206]. It is similar to the Mann-Whitney U test, since it compares medians. However, it allows to compare more than two groups.

6.2.6 Intra and Inter-observer Variability

The measurements by the same observer in different, but close occasions (intra-observer variability) by two independent observers (inter-observer variability) can be evaluated calculating the limits of agreement using Bland-Altman's plots. In these plots the differences between the studied parameters are plotted against the mean values. The coefficient of variation (CV) can also be used to assess the pulse-by-pulse variability of a dataset.

6.3 Results

6.3.1 Distribution of Values

The histograms in Figure 6.1 show the distribution of SWTT, LVET and RWTT for the GroupC and GroupH. As, it can be seen in Figure 6.1(a), SWTT occurs later for the hypertensive group (GroupH), where arrival times assume the mean value of 219.99 ± 51.03 *ms*, while for GroupC, this value is 144.12 ± 50.46 *ms*. It is visible the SWTT dependency from RWTT values, by the overlap of early reflected waves that coming before SWTT in GroupH. Concerning the LVET, its arrival time is quite similar for both groups: 313.52 ± 55.00 *ms* for the GroupH and 291.84 ± 35.28 *ms* for the GroupC (see Figure 6.1(b)). The RWTT occurs earlier for the GroupH (118.94 ± 39.95 *ms*), in opposition to the GroupC (200.46 ± 34.19 *ms*), as it is shown in Figure 6.1(c).

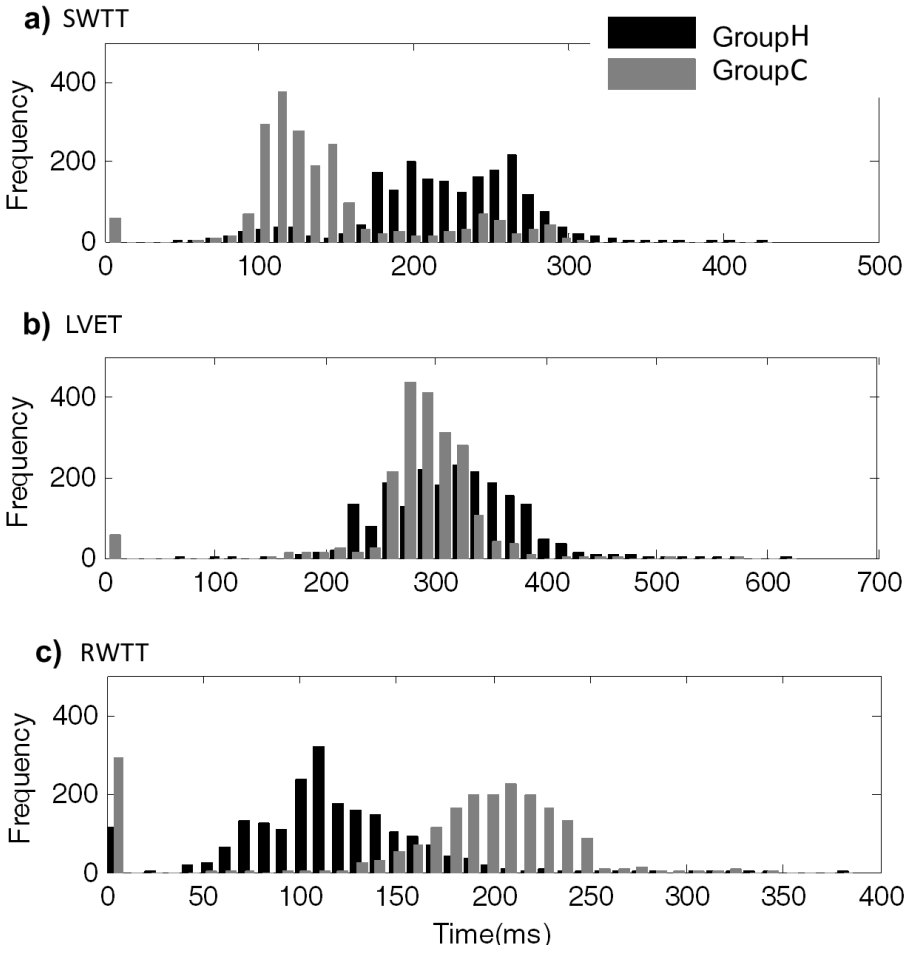

Figure 6.1: Arrival time frequency distribution for SWTT, LVET and RWTT.

When the algorithm is not able to identify the prominent points due to morphological artefacts, or algorithm errors, a null timing is assumed and, consequently, the first histogram bin is incremented. These events are more frequent for the RWTT analysis, in healthy subjects (GroupC), as it can be seen in Figure 6.1(c).

The magnitude of AIx was also computed for each group, as it is shown in Figure 6.2. The positive AIx values verified for GroupH are visible in black bars, and the negative values verified for the GroupC are visible in grey bars. For the GroupD, the positive values are also predominant, as it is visible in the white bar histogram distribution.

The RMSSD values intend to assess the variability associated to each one of the prominent points, in time and amplitude. RMSSD values for GroupH (black) and for GroupC (grey) are presented in Figure 6.3. For SWTT, it can be observed the presence of higher successive differences ($RMSSD_{SPt}$) for the GroupH. The similarities in both, $RMSSD_{DWt}$ and $RMSSD_{DWa}$ parameters reveal the low predictive value of DW in this dataset. Concerning the inflection wave point time, subjects from GroupH show a significant increase for the $RMSSD_{Pit}$ distribution.

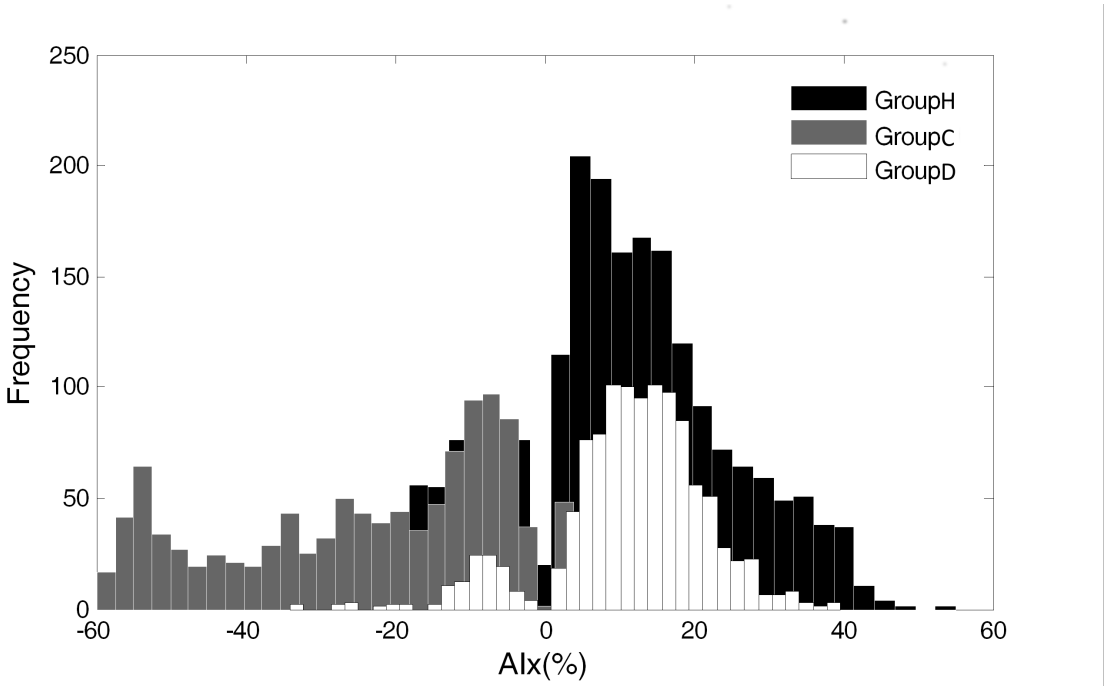

Figure 6.2: AIx values distribution for GroupC, GroupD and GroupH.

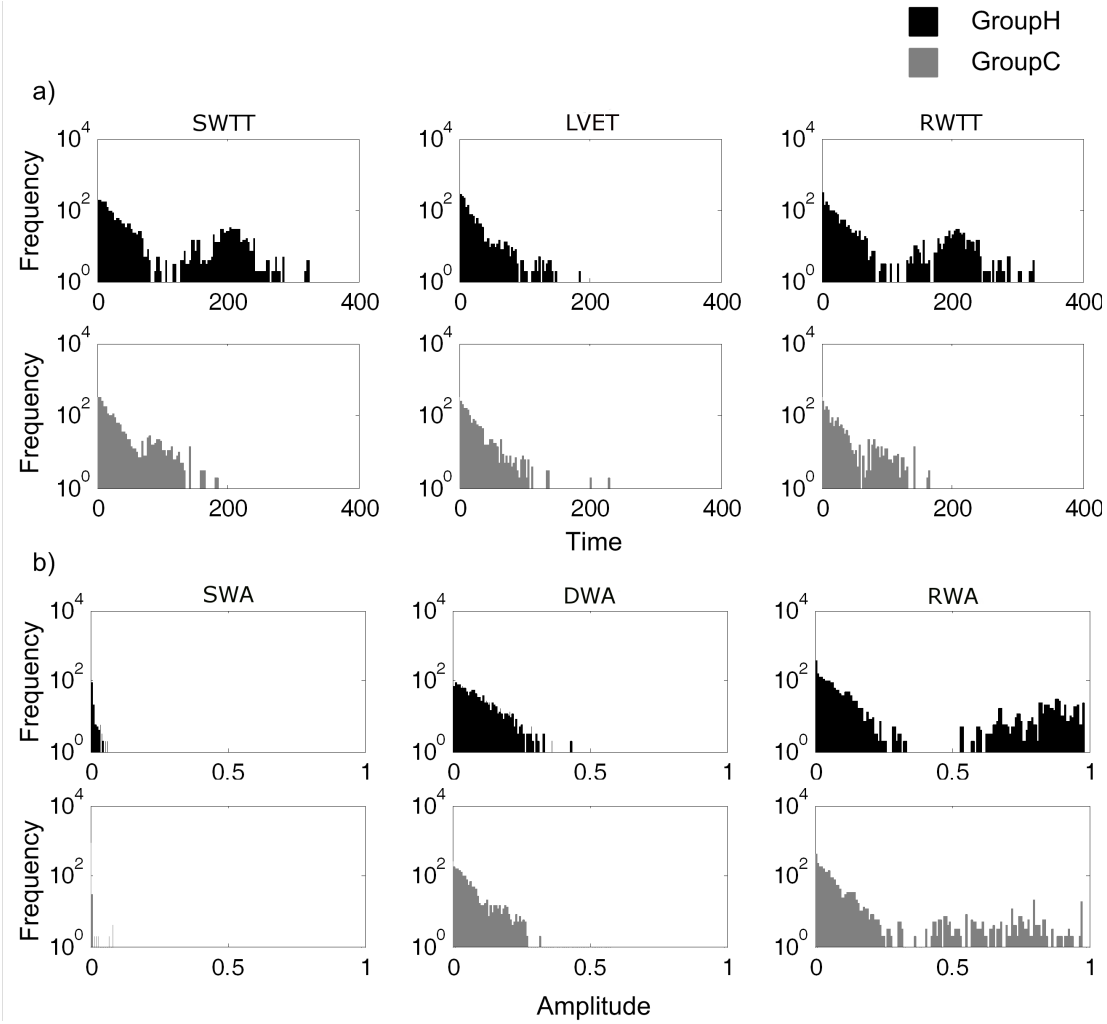

Figure 6.3: RMSSD successive differences measured for the SW, DW and RW: (a) time information (SWTT, LVET and RWTT); and (b) amplitude information (SWA, DWA, RWA).

6.3.2 Intra-arterial *versus* Non-invasive Waveforms

Data from GroupA were collected before and after a stent implementation, through invasive and non-invasive methodologies. The comparison between methods was performed through visual inspection.

The waveform comparison is shown in Figure 6.4 for a small segment ($3-4$ s), for both methods. The morphological changes that occur before and after the stent implementation are clearly visible, mainly for the Pi parameterizations (blue and red circles). The waveform differences were detected by the invasive device, but also by the non-invasive prototype under test. The time parametrizations obtained using the non-invasive prototype are shown in Figure 6.5. Before

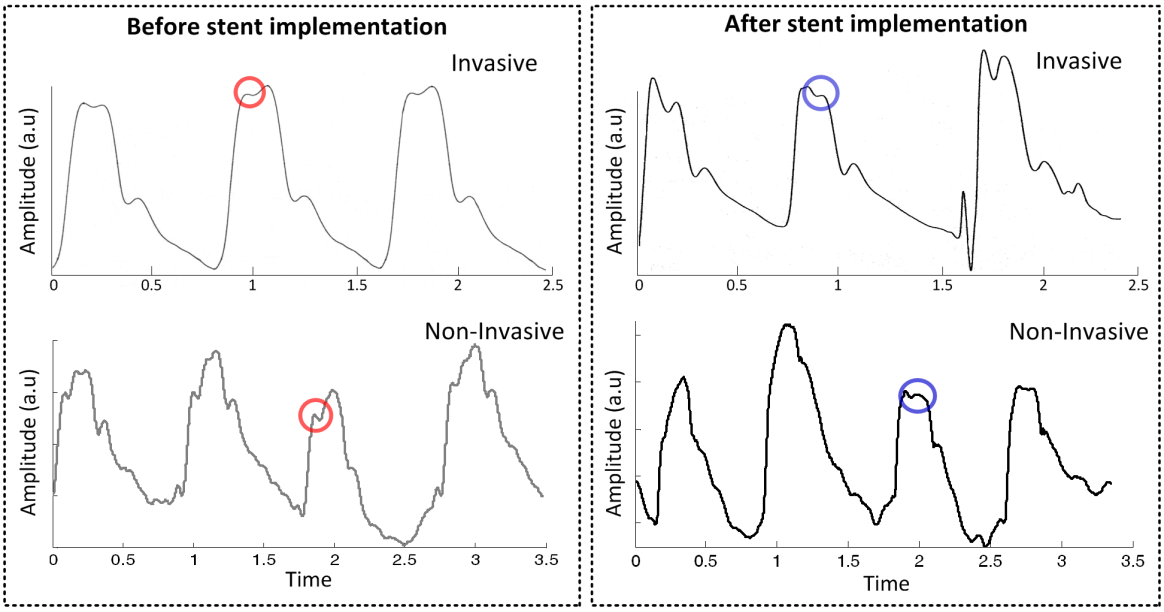

Figure 6.4: Set of pulses detected invasively and non-invasively after and before the stent implementation. The Pi arrival points for each method are represented by red and blue circles, for pulses collected before and after angioplasty, respectively.

the cardiac catheterization, RWTT occurs earlier, while SWTT appears later, whereas after the angioplasty, SWTT occurs prior to the RWTT. As expected, small changes in LVET values were registered, since were not verified known cardiac valve complications in this group of subjects.

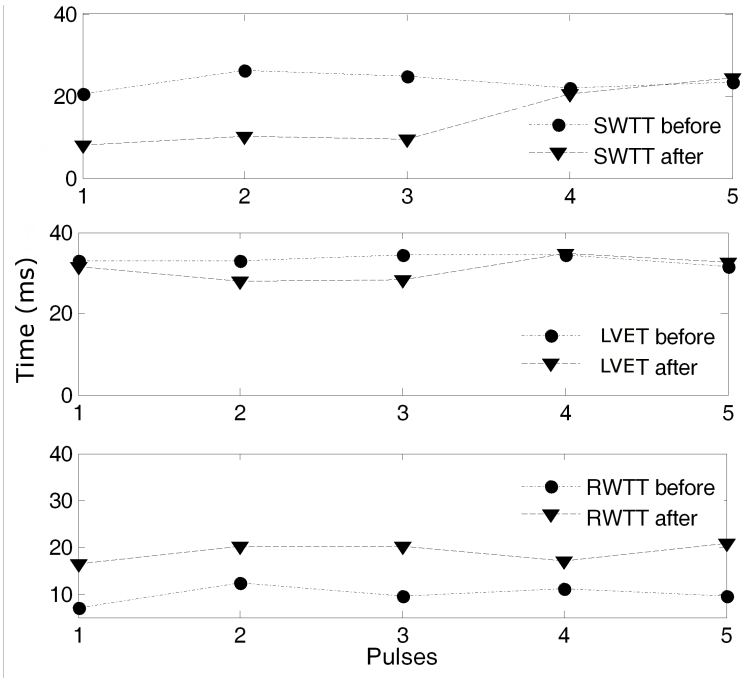

Figure 6.5: RWTT, SWTT and LVET comparison, before and after the angioplasty intervention.

6.3.3 Normality Assessment

The Kolmogorov-Smirov test was applied to the GroupR, for each trial, to assess the normality of the distribution. Each category in Trial 1 (month, carotid artery and sessions) and Trial 2 (operator, carotid artery and sessions) was analysed. It was observed that most of the variables were non-normally distributed. The non-normal distributions were expected, since they are quite common in large datasets, such as those used in this study. These results suggest that non-parametric approaches should be used. Few parameters presented normal-distributions in the Trial 2, namely the RWA values for the subject 4; SWTT, DWA and RWTT values for the subject 9; DWA values for the subject 10 and, RWA values for the subject 11. Other punctual situations were verified in the analysis of Trial 1, such as described by Borba, J. [184].

6.3.4 Significance of Differences

6.3.4.1 Categorization by Month

The monthly categorization was performed for the GroupR (Trial 1). Results are presented in Table 6.3. The Mann-Whitney U non-parametric test was used to compare the variables. It can be verified the existence of differences in the SWTT analysis between month 1 and month 2. Month 1 presents the lower SWTT average value (155.63 ± 58.76 *ms versus* 177.74 ± 64.26 *ms*). Average differences between months of about 7 *ms* and 2.5 *ms* were observed for the LVET and RWTT,

Table 6.3: Descriptive data categorization by month. The *p-value* indicates the Mann-Whitney test significance value.

Parameter	**Total**	**Month 1**	**Month 2**	***p-value***
SWTT (ms)	166.27±62.44	155.63±58.76	177.74±64.26	$< .05$
RWTT (ms)	167.70±37.36	168.90±40.12	166.41±4.09	$< .05$
LVET (ms)	289.47±52.27	285.98±61.72	293.23±39.31	$< .05$
SWA (a.u)[a]	0.99±0.01	0.99±0.01	0.99±0.01	$< .05$
RWA (a.u)[a]	0.87±0.08	0.86±0.09	0.88±0.07	$< .05$
DWA (a.u)[a]	0.71±0.14	0.70±0.15	0.72±0.13	$< .05$
AIx (%)	−1.16±15.05	−2.67±15.94	0.48±13.84	$< .05$

[a] Arbitrary Amplitude Units

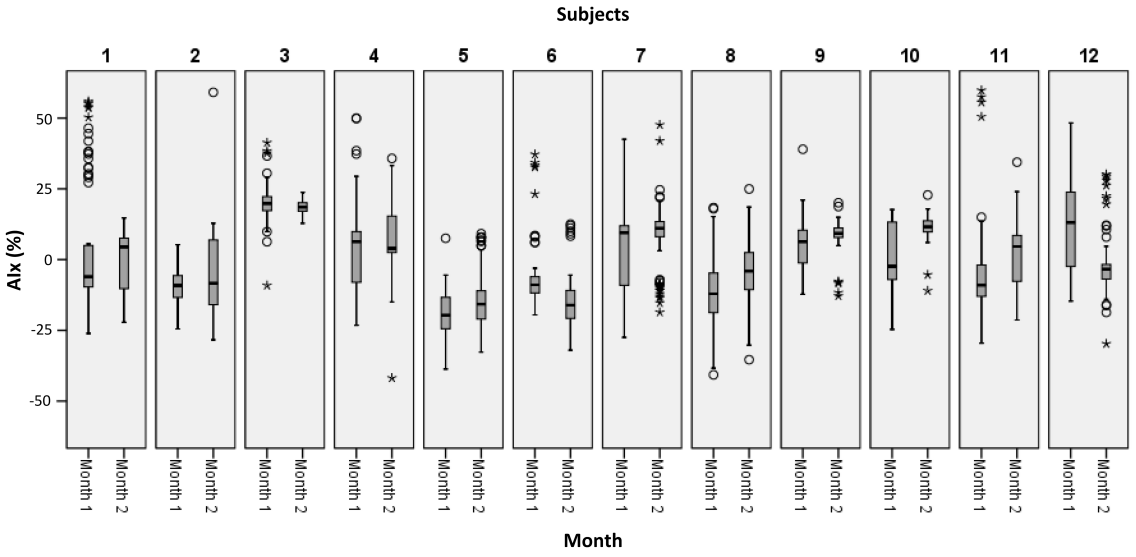

Figure 6.6: Boxplot of the AIx categorized by subject and sub-categorized by month.

respectively. AIx values are lower for month 1, comparing to month 2 (-2.67 ± 15.94 % *versus* 0.48 ± 13.84 %). According to the Mann-Whitney test, the verified differences are significant for all of the variables.

A categorization by subject and sub-categorized by month was performed and, the results concerning the AIx analysis are presented in Figure 6.6. Mann-Whitney U test demonstrated that only the subjects 1 and 4 do not present significant differences.

SWA was removed from subsequent statistical analysis, since the pulse amplitude was previously normalized in reference to this parameter during the pre-processing.

6.3.4.2 Categorization by Operator

The categorization by operator allowed the consistency determination for all the measurements, with the three operators. The subjects within the GroupR (Trial 2) were studied, and the descriptive data categorized by operator are shown in Table 6.4. The Kruscal-Wallis test was used, since it allows the comparison between more than 2 groups. The existence of significant differences for most of the variables was verified. The only exception was the RWTT ($p = 0.36$).

Table 6.4: Descriptive data categorized by operator. The *p-value* indicates the Kruscal-Wallis test significance value.

Parameter	**Total**	**Operator1**	**Operator 2**	**Operator 3**	***p-value***
SWTT (ms)	180.51±62.77	175.11±60.81	180.45±65.93	186.09±61.16	$< .01$
RWTT (ms)	157.25±42.67	156.94±44.06	156.45±40.68	158.35±43.08	0.36
LVET (ms)	286.63±40.95	281.06±40.01	288.60±44.14	290.43±38.03	$< .01$
RWA (a.u) [a]	0.85±0.11	0.86±0.12	0.85±0.10	0.84±0.11	$< .01$
DWA (a.u)[a]	0.76±0.13	0.76±0.12	0.77±0.12	0.76±0.13	$< .01$
AIx (%)	2.63±18.45	2.34±18.23	2.41±18.05	3.15±9.04	$< .01$

[a] Arbitrary Amplitude Units

6.3.4.3 Categorization for the Left and Right Carotid Arteries

Table 6.5 shows the observed differences for the left and right carotid artery, considering all operators. Data show for the RWTT parameter, that there was no significant differences between left and right measurements. It is also visible that the left carotid artery presents lower differences, for the SWTT, RWTT, and AIx analysis, comparatively to the right carotid artery. A detailed analysis, sub-categorized by operator, concerning this topic, is presented in Appendix E.1. Several statistical differences can be observed for the majority of parameters. However, there are exceptions, namely in the analysis of the RWA differences, with $p = 0.06$, $p < 0.05$ and $p = 0.51$, for the operators 1, 2 and 3, respectively. Non-statistical differences were also verified in the analysis of LVET (operator 2), and DWA (operator 1).

Table 6.5: Kruskal test significance for the left and right measurements.

Parameter	***p-value* (Left)**	***p-value* (Right)**
SWTT	0.73	$< .01$
LVET	$< .01$	$< .01$
DWA	$< .01$	$< .01$
RWTT	0.22	0.16
RWA	$< .01$	$< .01$
AIx	0.12	$< .01$

6.3.4.4 Categorization by Subject

A statistical analysis for each subject, sub-categorized by operator and carotid artery was performed to assess the differences among subjects. The Kruskal-Wallis test was conducted to assess these dissimilarities, and the results presented in Tables 6.6 and 6.7.

Figures 6.7 and 6.8 show the AIx median values for each subject, sub-categorized by operator, for the right and left carotid arteries, respectively. The analysis of the subjects 3, 4 and 19 evidences a low number of pulses in the dataset for one of the operators (not the same for all). So, these subjects were not included in the subsequent analysis due the lower number of pulses, comparatively to the other subjects. Discrepancies in AIx values (*i.e.* positive and negative values) are visible for the subjects 1 and 8 (right artery) and for the subjects 1, 6, 15 and 16 (left artery).

Other data were plotted, considering each subject and operators individually, being presented in Appendix E.2. For all parameters, it can be observed that dicrotic features (DWA and LVET) have lower significant differences, comparatively to the other parameters. However, for the majority of the parameters, significant differences were found.

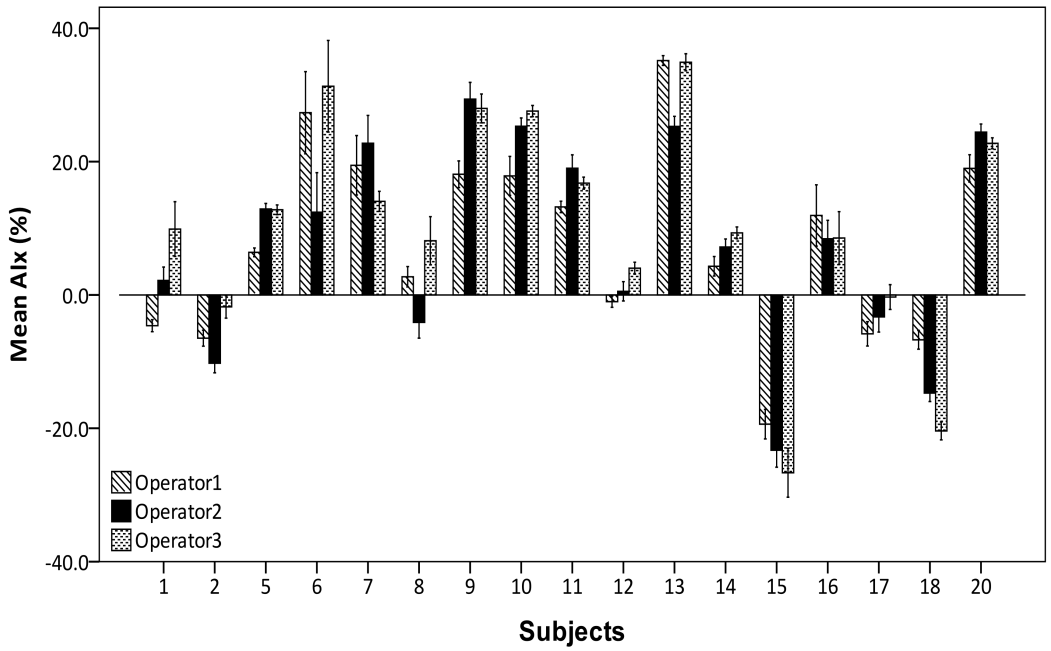

Figure 6.7: AIx values (right carotid artery) obtained for each subject and operator.

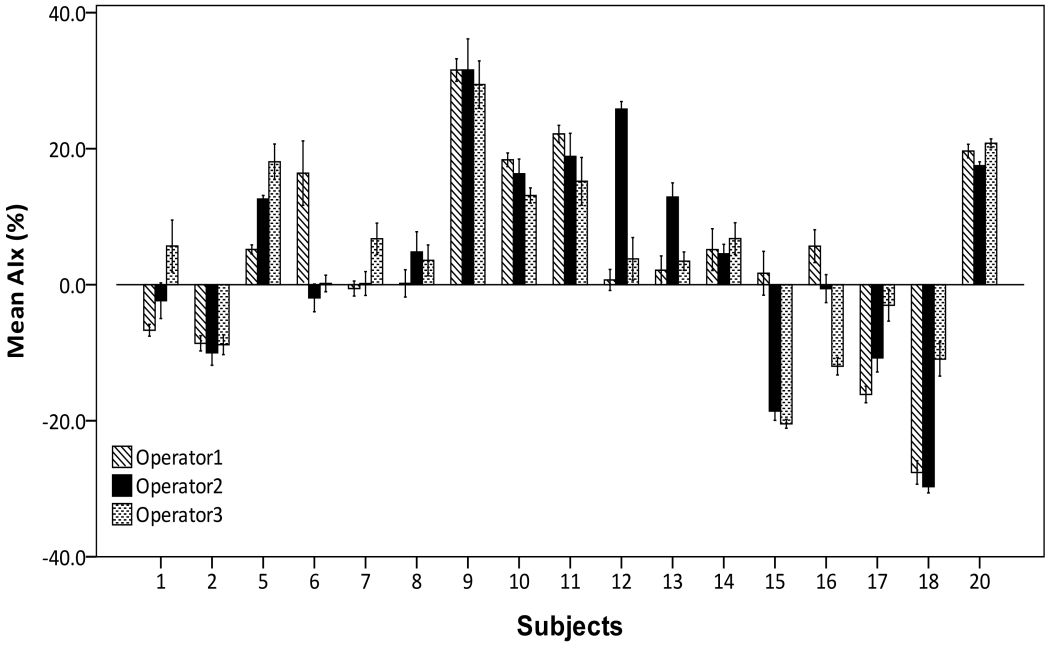

Figure 6.8: AIx values (left carotid artery) obtained for each subject and operator.

Table 6.6: Descriptive data categorization by subject for the left carotid artery.

Parameter	Subject1	Subject2	Subject5	Subject6	Subject7	Subject8	Subject9	Subject10	Subject11
SWTT (ms)	144.27±48.33**	129.52±41.07**	249.69±32.72**	160.03±54.07**	179.68±53.92*	229.86±13.16**	272.30±15.20	211.29±1.50**	263.31±31.25**
RWTT (ms)	165.48±19.89**	172.40±19.81**	162.14±24.94**	154.36±41.29**	167.20±20.31**	197.53±62.91*	128.63±17.08**	113.19±25.98**	153.17±19.78**
LVET (ms)	238.85±55.12**	268.32±24.35**	324.33±17.22**	261.00±41.74**	279.38±26.12	301.44±36.14**	334.94±18.30	291.61±24.42**	333.30±28.50
RWA (a.u)[a]	0.92±0.03	0.88±0.05**	0.89±0.06**	0.88±0.14**	0.92±0.05**	0.89±0.06	0.69±0.05	0.83±0.06**	0.79±0.06**
DWA (a.u)[a]	0.87±0.08**	0.70±0.08*	0.79±0.04**	0.80±0.09**	0.78±0.06**	0.85±0.06**	0.90±0.06**	0.77±0.07**	0.86±0.06**
AIx (%)	−4.67±7.55**	−9.10±8.85**	10.09±6.79**	4.31±17.90**	1.29±9.68**	2.27±11.84**	31.01±5.17	15.73±8.71**	20.10±8.01**

Parameter	Subject12	Subject13	Subject14	Subject15	Subject16	Subject17	Subject18	Subject20
SWTT (ms)	188.49±43.23**	196.92±63.11**	176.25±54.09**	115.04±6.44**	163.23±51.54**	122.47±47.88**	121.56±45.60**	234.10±20.92**
RWTT (ms)	170.72±53.09**	157.64±18.97**	134.67±27.22**	172.76±28.92**	186.97±42.16**	178.75±36.06**	186.49±15.54*	117.30±17.71**
LVET (ms)	274.50±28.76**	311.15±26.32**	272.75±29.72	263.92±33.43**	296.45±43.03**	289.19±49.23**	283.36±17.53**	307.32±18.88**
RWA (a.u)[a]	0.91±0.10**	0.88±0.05**	0.90±0.06**	0.82±0.09**	0.89±0.07**	0.86±0.07**	0.74±0.07**	0.80±0.05**
DWA (a.u) [a]	0.85±0.07**	0.71±0.10**	0.82±0.08*	0.58±0.12**	0.80±0.09**	0.70±0.11**	0.65±0.10*	0.75±0.07**
AIx (%)	4.20±12.44**	6.14±11.00**	5.33±10.48	−15.32±12.39**	−2.88±12.19**	−11.19±11.43**	−22.05±14.80*	19.47±4.76**

Significant levels: ** $p < 0.01$, *$p < 0.05$

[a] Arbitrary Amplitude Units

Table 6.7: Descriptive data categorization by subject for the right carotid artery.

Parameter	Subject1	Subject2	Subject5	Subject6	Subject7	Subject8	Subject9	Subject10	Subject11
SWTT (ms)	189.86±66.48**	132.46±48.07**	255.82±26.71**	159.89±45.51**	220.09±19.72**	216.83±29.37**	265.69±16.76**	231.47±26.49**	270.68±26.25**
RWTT (ms)	171.53±19.22**	165.06±23.36	167.80±18.31**	112.20±59.43**	129.85±37.63**	182.59±57.33**	131.42±27.04*	110.15±9.81**	161.07±17.57**
LVET (ms)	298.78±41.91**	266.88±27.10**	321.55±13.85**	237.68±50.89**	279.25±19.30**	284.66±19.07**	328.84±18.97**	303.89±23.01**	348.76±24.56**
RWA (a.u)[a]	0.92±0.05**	0.90±0.06**	0.89±0.05**	0.74±0.24**	0.82±0.10**	0.91±0.08**	0.73±0.10**	0.75±0.07**	0.84±0.04**
DWA (a.u) [a]	0.79±0.10**	0.71±0.10**	0.82±0.05**	0.87±0.10	0.89±0.05**	0.84±0.05**	0.87±0.06**	0.84±0.06**	0.85±0.05**
AIx (%)	−4.67±7.55**	−6.39±9.41**	10.48±5.67**	23.04±26.44**	16.95±11.46**	3.09±11.97**	26.59±9.51**	25.36±6.89**	15.29±5.39**

Parameter	Subject12	Subject13	Subject14	Subject15	Subject16	Subject17	Subject18	Subject20
SWTT (ms)	195.23±38.41**	210.50±63.11**	194.29±39.96**	112.43±10.08**	198.18±52.28**	150.01±56.30**	103.96±26.03**	233.44±22.22**
RWTT (ms)	166.65±30.29**	87.83±18.23**	149.06±16.41**	191.12±33.49*	166.23±56.16*	165.25±34.07	161.68±31.44**	108.60±28.49**
LVET (ms)	280.91±21.64**	313.64±43.54**	276.59±22.97	295.21±27.43**	280.00±48.17**	281.03±52.41**	263.73±50.05**	305.52±16.33**
RWA (a.u)[a]	0.95±0.0.04**	0.68±0.06**	0.90±0.04**	0.77±0.17**	0.88±0.13*	0.92±0.06**	0.85±0.09**	0.77±0.07**
DWA (a.u) [a]	0.82±0.06**	0.84±0.07	0.78±0.06**	0.58±0.08**	0.90±0.05**	0.79±0.11*	0.68±0.13**	0.72±0.10**
AIx (%)	1.48±5.97**	31.60±7.38**	7.00±7.83**	−22.86±17.08**	9.51±15.51*	−2.52±10.00*	−14.21±9.81**	21.96±8.47**

Significant levels: ** $p < 0.01$, * $p < 0.05$

[a] Arbitrary Amplitude Units

6.3.5 Intra-observer Variability

For the calculation of the intra-observer reproducibility, each of the three sets of data records, performed by each operator, were analysed. So, for each operator 3 pairs of values ($n = 51$) were considered (session1 *versus* session2, session1 *versus* session3, session2 *versus* session3), for both arteries (left and right). Bland-Altman plots were analysed for all data and, the observed differences are presented in Tables 6.8 and 6.9, for the left and right carotid arteries, respectively.

Generally, the verified differences were lower for the observer 3, comparatively to the others, which is probably due to its larger experience in performing pulse wave measurements. These data are consistent with the significant differences observed in Section 6.3.4.3. The results are similar (or better) to the other data reported in literature (see Table 2.2). In the AIx analysis, it was observed a

Table 6.8: Intra-observer mean differences $\pm$ 2 *SD* derived from the Bland-Altman plots ($n = 51$, obtained for the left carotid artery.

Parameter	**Operator1**	**Operator2**	**Operator3**
SWTT(ms)	5.74±17.32	5.41±15.12	1.71±29.08
RWTT (ms)	−2.4±18.54	−5.09±10.85	2.74±15.50
LVET (ms)	3.44±13.05	6.59±22.36	1.53±21.84
RWA (%) [a]	1.02±4.87	−0.90±3.13	−0.01±3.59
DWA (%) [a]	0.96±4.41	0.24±4.51	0.37±4.85
AIx (%)	−0.36±5.16	0.96±3.26	0.52±3.58

[a] In reference to the normalized amplitude

Table 6.9: Intra-observer mean differences $\pm$ 2 *SD* derived from the Bland-Altman plots ($n = 51$, obtained for the right carotid artery.

Parameter	**Operator1**	**Operator2**	**Operator3**
SWTT(ms)	4.18±19.74	−4.18±30.41	0.61±25.46
RWTT (ms)	−2.27±17.67	−8.97±23.80	2.70±17.24
LVET (ms)	0.19±15.73	−4.32±22.67	3.63±25.36
RWA (%) [a]	−0.16±3.29	−1.18±8.06	−0.92±7.63
DWA (%) [a]	0.59±4.38	0.90±5.34	−0.31±4.31
AIx (%)	0.02±6.01	2.18±6.70	0.90±7.21

[a] In reference to the normalized amplitude

difference in the intra-variability < 1 %, for the majority of cases (excluding the 2.18 % variability value, observed by the operator 2, at right carotid artery). The significance of the differences was also assessed using the Kruskal-Wallis test. The results are presented in Appendix E.3. These results confirmed that the operator 3 presented a better performance, since the measurements have lower significant differences among sessions.

Frimodt-Moller *et al.* [133] reported an RWTT variance of -6.9 ± 52.7 *ms*, -2.7 ± 32.8 *ms*, -3.2 ± 33.9 *ms*, using three different sets of measurements, similar to the values verified in this study, 2.27 ± 17.67 *ms*, 8.97 ± 23.80 *ms*, 2.70 ± 17.24 *ms*, for the operator 1, 2 and 3, respectively. The other timing variables (SWTT, RWTT) also evidence low differences, for the same order of values. RWA and DWA mean differences were about of 1 % for all measurements.

6.3.6 Inter-observer Variability

The calculation of the inter-observer variation was based on the average of 153 double records performed by all the operators (each pair of values comprises two of the three operator measurements). Results are presented in Table 6.10. Higher differences were observed for the right artery, comparatively to the left artery. In this case, an AIx value of -2.31 ± 7.29 % and a SWTT value of -12.94 ± 31.46 *ms* was observed, which are higher than the left measurements, 0.94 ± 7.52 % and -2.96 ± 22.67 *ms*, respectively. Probably, these values are related with the implemented acquisition protocol and operators positioning.

Siebenhofer *et al.* [134] reported an AIx mean difference of 0.4 ± 6.4 %, while Crilly *et al.* [135] reported 1.0 ± 3.9 % mean value and, Frimodt-Moller *et al.* [133] state up 0.9 ± 15.8 %, similar to the values verified in this study (0.94 ± 7.52 %). It can be also observed that the lowest inter-observer variability was recorded for the RWTT, 0.83 ± 27.10 *ms*, which represents an excellent value comparing to the Frimodt-Moller *et al.* [133] observed value of 1.9 ± 30.8 *ms*. RWA and DWA mean differences presented good variability range (< 3 %) for all measurements.

Table 6.10: Inter-observer mean differences ± 2 *SD* derived from the Bland-Altman plots ($n = 153$).

Parameter	Left	Right
SWTT(ms)	−2.96±22.67	−12.94±31.46
RWTT (ms)	−0.83±27.10	−1.08±23.49
LVET (ms)	−5.83±26.06	−9.57±23.03
RWA (%) [a]	−0.25±7.58	2.46±7.21
DWA (%) [a]	0.67±7.39	−1.07±6.86
AIx (%)	0.94±7.52	−2.31±7.29

[a] In reference to the normalized amplitude

6.3.7 Pulse-by-pulse Variability

There are no previous studies concerning the pulse-by-pulse variability for the APW features. In this study, the CV was assessed for all pulses in the dataset (comprising session1, session2 and session3) measured by each operator. Results are presented in Figure 6.9. Variability values were very good for the LVET, DWA, RWA (CV< 10 %), but unsatisfactory for the AIx (28.07 %, 21.51 % and 24.85 % for the operators 1, 2 and 3, respectively). And, the SWTT and RWTT presented satisfactory results (10 %<CV>20 %). Only the AIx has unsatisfactory pulse-by-pulse variability (in spite of borderline satisfactory) and, therefore careful standardization of this parameter is highly required. Thus, AIx measurement should be reserved mainly to highly trained operators. The other

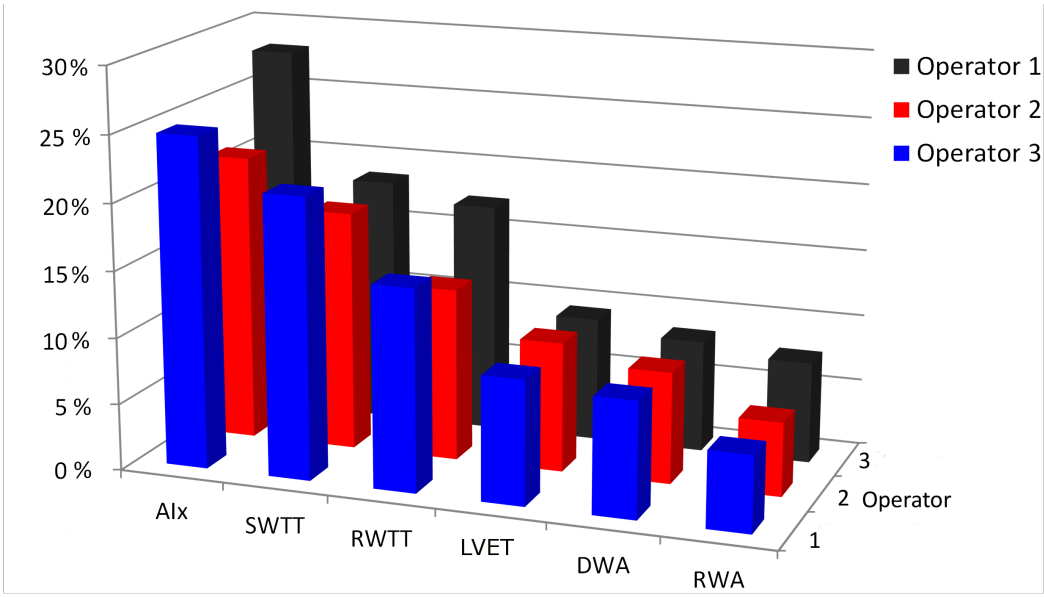

Figure 6.9: AIx values (right carotid artery) obtained for each subject and operator.

parameters give good indicators concerning their potential use for cardiovascular risk assessment.

6.4 Conclusion

The validation tests demonstrated the feasibility of creating a database for research and clinical practice. By indirect comparisons and internal consistency tests, it has been found that the database created has enough quality for research purposes, presenting a wide range of parameter values.

In general, non-invasive measurements are less accurate, when compared with invasive trials, mainly due to the dependency of the measurement conditions and operators. However, the non-invasive APW module is capable of effectively detecting physiological modifications, before and after a stent implementation. On an angioplasty procedure was demonstrated the usefulness of this technology to ease the early identification of cardiac problems, when used in clinical practice.

The intra-observer ”limits of agreement” for the studied parameters, SWTT, RWTT, LVET, RWA, DWA and AIx, are good markers comparatively to the values available on the literature. Additionally, it was demonstrated that operators can be easily trained in the use of this prototype, since different operators with only one week of training produced results quite reasonable. These results have important implications concerning the use of this prototype for the clinical assessment of patients and suggest that APW indices are reproducible and repeatable when measured by different observers at repeated occasions.

A set of more well defined and demanding acquisition protocols could and should be implemented, in order to minimize even further any possible external influences. Finally, it is possible to suggest the application of this technique at a large scale, randomised clinical trials to address important clinical questions, such as applications to anti-hypertensive therapies and cardiovascular interventions.

7

Concluding Remarks

7.1 General Conclusions

The main motivation of this work was to provide a valid contribution for the assessment of the cardiovascular condition. In order to accomplish this goal, the adopted methodology consisted of: i) the development of a multichannel platform capable to add up several technologies and methodologies in order to provide a refined cardiovascular assessment; ii) the development of data-mining tools, incorporating personalized machine learning algorithms capable to deal with a wide set of APW features; iii) the construction and population of a cardiac database essential to the system validation.

The multichannel platform allows to integrate several technologies and methodologies used in the traditional clinical path of cardiovascular patients (*e.g.*, oximetry, ECG, PWV). The main instrumental focus of this work remained on a module dedicated to APW monitoring, being the PZ sensors the basis of the developed probe. The performance of the probe was assessed using a dedicated purpose test bench, capable of generating arbitrary shaped pressure waves over a confined artery-like flexible tube filled with an appropriate liquid. The performed tests in this environment comprised an important role along the development of the instrumentation module and probe, since they allowed the optimization of the signal conditioning circuit, and the choice of the optimal probe design. The results demonstrated that the PZ probe performed very well in a number of fundamental accuracy tests on elastic tubes. The maximum RMSE introduced by the electronic circuit itself is 0.33 % and the total average error (including the mechanical interface contribution) does not go higher than to 1.80 %.

Concerning the algorithmic tool, the implemented processing tasks included algorithms for baseline fluctuation elimination, pulse segmentation, establishment of the adequate bad pulse flagging criteria and the development of a waveform delineator algorithm for the identification and quantification of the most relevant APW features. Most of APW analysis frameworks individually process each hemodynamic parameter and do not evaluate the inter-dependencies between features. With this work, an extensive pool of parameters was established: time/amplitude intervals, some ratios and indices concerning APW parametrizations and pulse variance measures. The extraction of the appropriate features was essential for the development of the decision support system platform.

The main advantage of the implementation of innovative decision support system algorithms is the possibility to deal with a wide set of APW parameters in a broader variety of operating situations, even in the case of missing values. It is possible to obtain important indices concerning the arterial stiffness that can be used as estimates of the risk associated to the APW morphology. Classification techniques have shown their usefulness in assessing CVD risk. The J48 or Random Forest algorithms revealed to be appropriate options, that can be used by the clinicians. Neural networks have also demonstrated good performance In general, for the tested classifiers, higher accuracy values were obtained (above 95 %), as well as AUC values comparable to other models described in the general cardiology literature. It was also demonstrated the potential use of multiple classifier methodologies in producing more trustworthy diagnostic outputs in comparison with single classifier analysis. Clustering procedures also revealed an important role, since they allowed the premature determination of different CVD risk groups. These methods are particularly interesting since can avoid the requirements of the classification procedures which often require costly labelling of a large set of parameters, such as biochemical analysis used by traditional risk scores. One of limitation of the proposed analysis methodology is its high dependence on the database subjects. This means that for a successful modelling, data must be representative, *i.e.* they must cover all of the possible physical situations.

The validation tests were essential in the assessment of the clinical use of the APW module. A performed case study in subjects with carotid stenosis was accomplished, enabling the detection of APW modifications before and after the angioplasty procedure. The inter and intra-observer limits of agreement for SWTT, RWTT, LVET, RWA, DWA and AIx, are good markers comparatively to the values available on literature, even when APW acquisition is undertaken by less trained operators.

7.1.1 Future Research Lines

A long-term therapeutic trial for investigate the role of arterial stiffness, systolic blood pressure and biochemical parameters, such as RAAS, should be planned in order to identify the most important risk predictors associated with the development of CVD and correlate them with APW modifications.

Additionally, it would be interesting to study the relationship between transthoracic echocardiography and the APW parameters. The relationship between the increased load imposed on the left ventricle by wave reflections and arterial stiffness parameters, such as AIx needs to be widely explored [113, 210]. An interesting outcome at this stage could be the relationship between the echocardiography parameters, such as, the increase in Left Ventricular Mass Index (LVMI) and cardiac output, with the APW features.

Other interesting application could be the use of this platform (hardware and its algorithmic tools) during the anaesthesia monitoring in surgery environment. Currently, the measurement of the APW features, crucial for the patient condition monitoring, is accomplished through invasive methods. This procedure is important for the cross-correlation with other routine parameters in order to assess the volemic state of the patient and the nociception balance. The availability of non-invasive sensors will enable the full arterial pulse profile measurement, avoiding invasive probes. Moreover, this approach is particularly suitable to use in intensive care patients with hemodynamic instability also for volemic characterization and analgesia dose adequacy. The outcome of this approach could be a set of APW parameters chosen to fully characterize the patient anesthesia/analgesia demand over time. A training period with both, invasive and non-invasive probes would be the basic requirement for the establishment of the most suitable classifiers (through data mining) to deal with the set of parameters.

The developed prototype must be validated against SphygmoCor and Complior devices. The evaluation is mandatory along with the demonstration of the PZ sensor, in large clinical and research units. The database established along this study comprises 246 subjects, but there is lack of data from unhealthy subjects and, data from healthy subjects in the $30-50$ age range.

In the field of machine learning algorithms some improvements can also be performed. The application of clustering techniques in a sample composed by patients with different CVD, can be relevant to assess the power of the clustering methodologies in partitioning the groups by different cardiovascular pathologies. Future improvements can also focus on the study of other machine learning algorithms, such as Support Vector Machines (SVMs) algorithms that can improve the accuracy, due to the known ability to deal with non-categorical features. SVMs are frequently referred to as the classification technique with higher overall accuracy [211].

Appendices

A Database Structure

Main menu includes several icons useful to perform common operations and a top menu bar that allows for the selection of the operations. Table A.1 presents a description of main menu shortcuts.

Table A.1: Cardiocheck main menu icons.

Icon	Description
Acquisition 	Opens *Acquisition* window where are available several settings: the operator and its institution, the place of acquisition, the patient, the hardware modules and the time settings.
Search 	Opens *Search* window that allows the search of patients, institutions, operators and hardware modules. There are some filtering options, that allows to adapt the search results to the desired conditions.
Add 	Opens a side menu that allow for the creation of a new patient, institution or operator.
Processing 	Opens *Processing* window where the acquired data can be viewed and processed.
Settings 	Opens a side menu, where a patient, institution or operator can be excluded or restored from the database. It is also possible the creation of a new database or, the choice of another database available in the computer. The access to this menu is limited to an administrator, since it requires a password.
Exit 	Closes all windows that might be opened and exits the Cardiochek interface.

The *Acquisition* menu is dedicated to the operator, patients, acquisition place. An overview of this window is presented in Figure A.1.

Search menu is intended to perform searches on patients, institutions and operators inserted into the Cardiocheck database. This menu is presented in Figure A.2. This process is essential to perform the data analysis for different subsets.

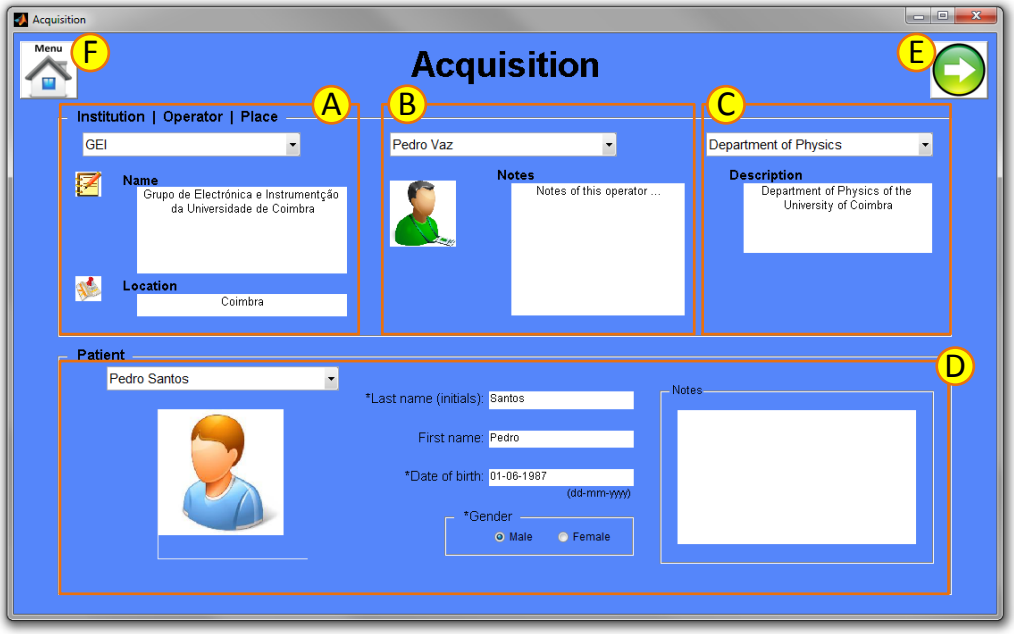

Figure A.1: Cardiocheck *Acquisition* window: institution and patient settings. This window has the following fields: (A) institution select pop-up; (B) operator select pop-up and a field for notes; (C) place of acquisition select pop-up; (D) patient select pop-up and the corresponding information and notes; (E) proceed to signals and hardware window; (F) main menu button.

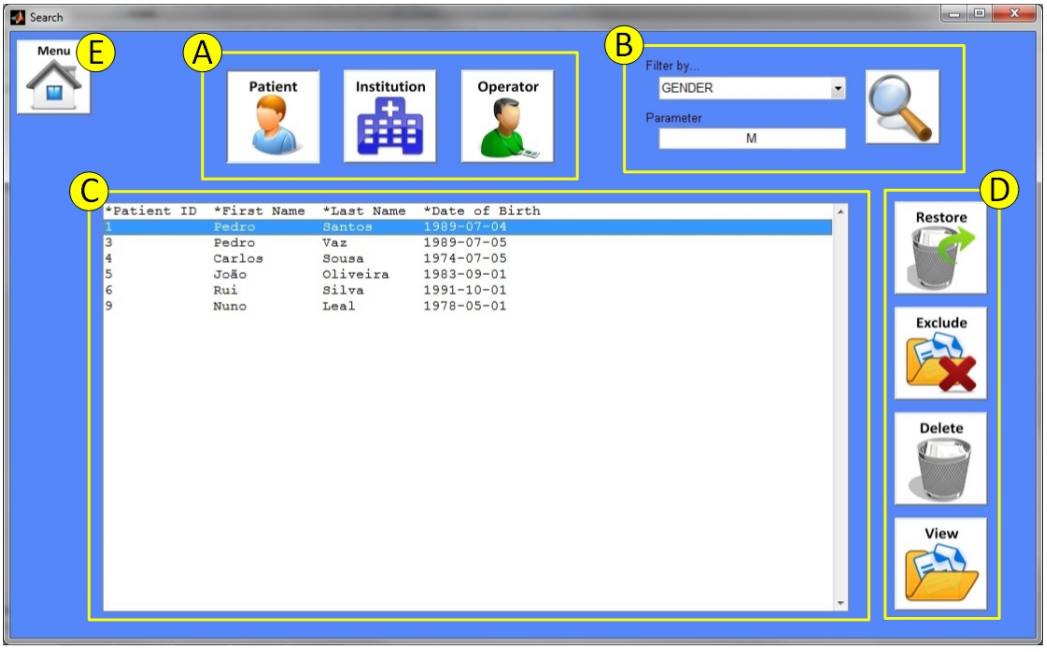

Figure A.2: Cardiocheck *Search* menu: simulation of a patient search with ”M” gender filtering condition. Fields in search window include: (A) entity selection; (B) filtering conditions; (C) search result list; (D) operations buttons; (E) main menu button.

Patient window has four tabs with diverse information: personal info, clinical history, family history and biochemical analysis. When inserting a new patient, it is possible to fill these fields, but the edition *a posteriori* is also allowed. Additionally, for each acquisition are also required information and institution and operator informations. This interface communicates with other window capable to deal with signals from the modules of the multichannel platform, namely with the algorithms described in Section 4.4.

B Simulink Configurations

The Master Block configurations (Table B.1) set important dsPIC options like PLL, oscillator mode or clock frequency.

Table B.1: dsPIC MASTER

Number of Instructions Per Second	50000000
Oscillator Mode	Quartz (XT - HS)
Clock frequency	8e6
PLL	Activate
Timer step	0.001
Error	0 %
PR1	49999

The UART Configuration block (Table B.2) set the parameters for the UART. Each UART has one Receive (Rx) pin and one Transmit pin (Tx). The following parameters are fixed: 1 stop bit, 8 bits data, no parity and no flow control. Information about the real baud rate obtained and the corresponding % error, comparatively to the selected baud rate is obtained from simulink. The % error must be near or lower than 3 % for the RS-232 serial transmission work properly [180].

Table B.2: UART 2 Configuration

Baud Rate (kb/s)	4608800
UxBRG	6
Real Baud	446429
Error	-3.12 %
	44.6429 Bytes/step

C WEKA Toolkit

Weka is a collection of machine learning algorithms for data mining tasks. The algorithms can either be applied directly to a dataset or called from Java code. Weka contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization. Weka's main graphical user interface, the *Explorer* menu, gives access to all its facilities using menu selection and form filling. It is illustrated in Figure C.1. There are six different panels that can be selected by the tabs at the top, corresponding to the various data mining tasks that Weka supports.

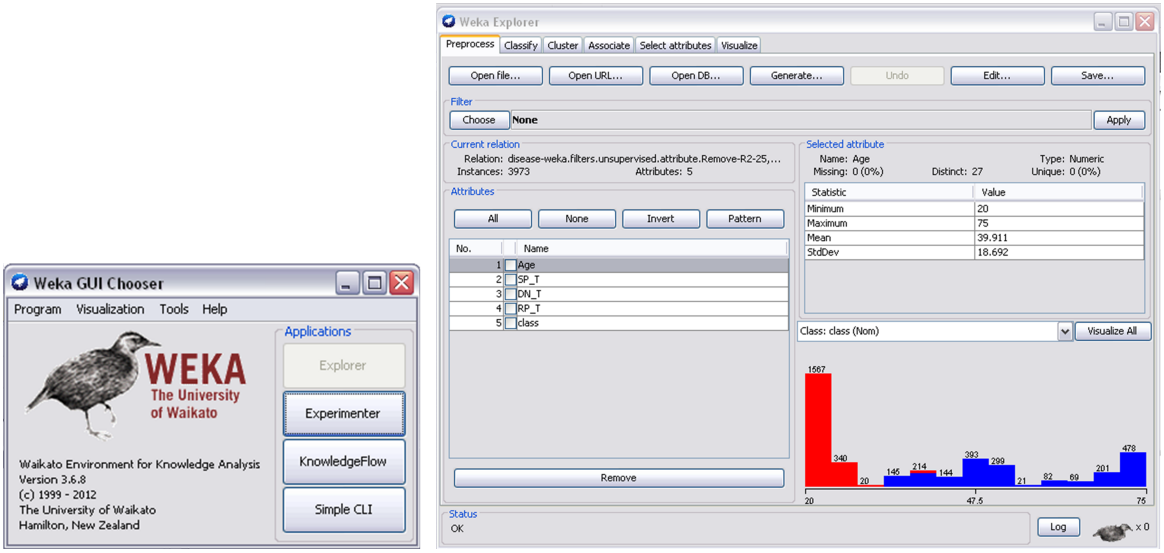

Figure C.1: The Weka Explorer: (a) choosing the *Explorer* interface and (b) visualization of the data (*Preprocess* panel).

In summary, the main functions of each one of tabs presented at the top of the *Explorer* menu are described below. The *Clustering* and *Classification* panels are shown in Figure C.2.

1. Preprocess - Choose the dataset and modify it in various ways.
2. Classify - Train learning schemes that perform classification or regression and evaluate them.
3. Cluster - Learn clusters for the dataset.
4. Associate - Learn association rules for the data and evaluate them.
5. Select attributes - Select the most relevant aspects in the dataset.
6. Visualize - View different two-dimensional plots of the data and interact with them.

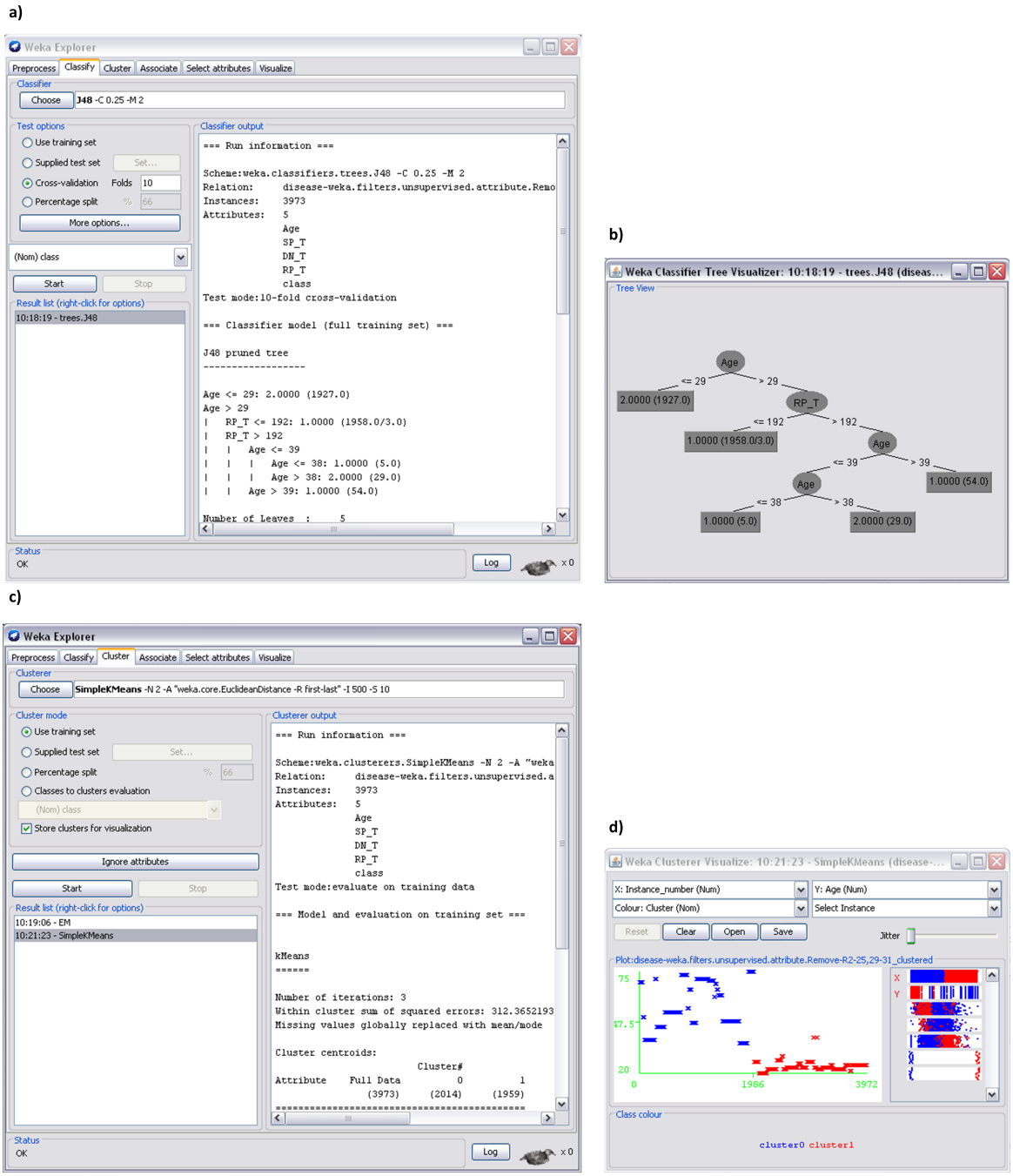

Figure C.2: Visualizing the result (a) of J48 classifier algorithm and, (b) its tree; (c) *k*-means clustering results, and (d) the cluster distributions.

The *Visualize* panel helps to visualize a dataset, not only the result of a classification or a clustering model, but the dataset itself. It displays a matrix of two-dimensional scatter plots for each pair of attributes. Figure C.3 shows the distributions for a studied dataset.

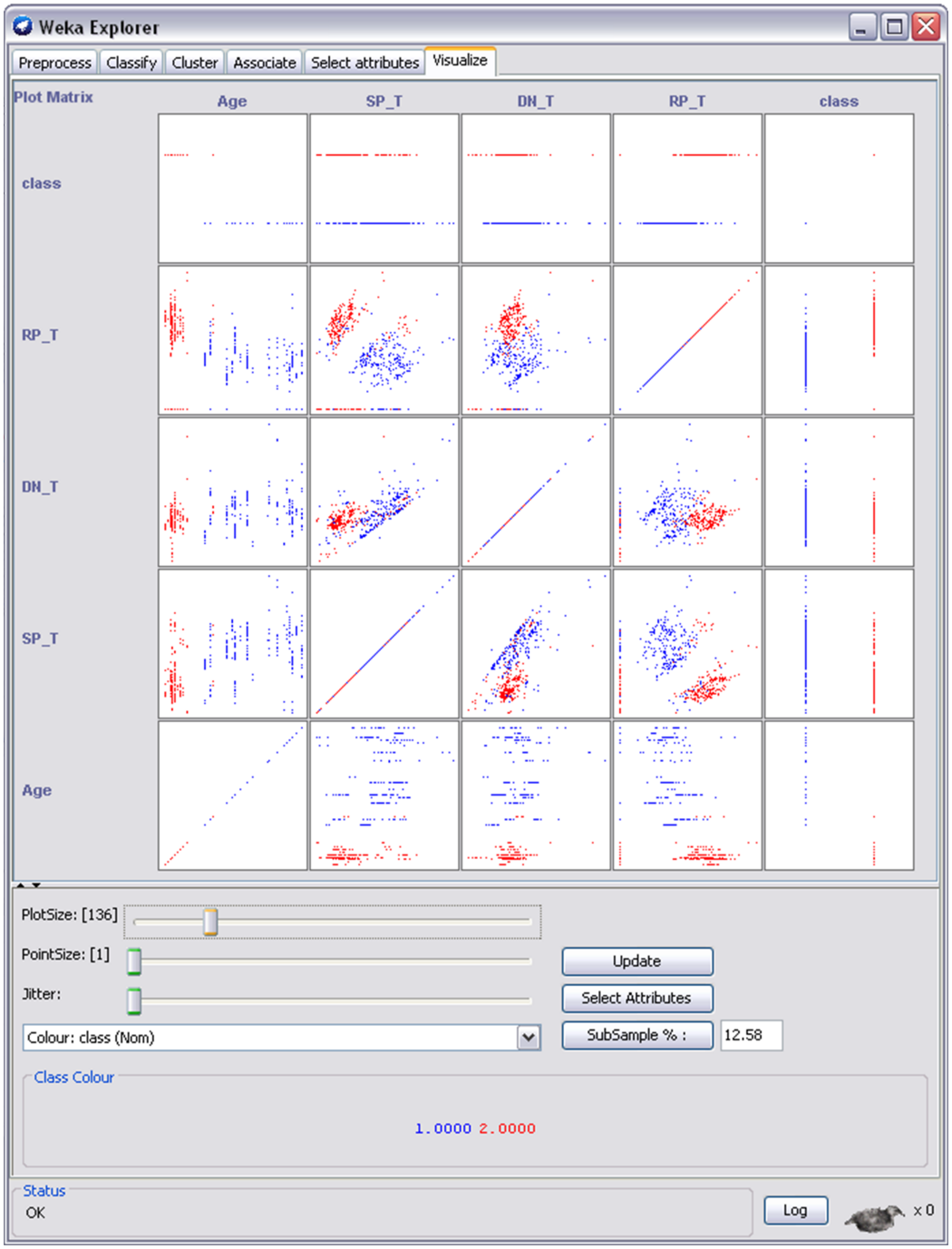

Figure C.3: *Visualize* menu.

D Scores

D.1 Framingham Risk Score (FRS)

Since 1948, the Framingham Heart Study, under the direction of the National Heart, Lung and Blood Institute (NHLBI), formerly known as the National Heart Institute (USA), has been committed to identifying the common factors or characteristics that contribute to cardiovascular disease (CVD). This study followed CVD development over a long period of time in several generations of participants. It was used the score for General Cardiovascular Risk (10-year risk), such as presented in Figure D.1.

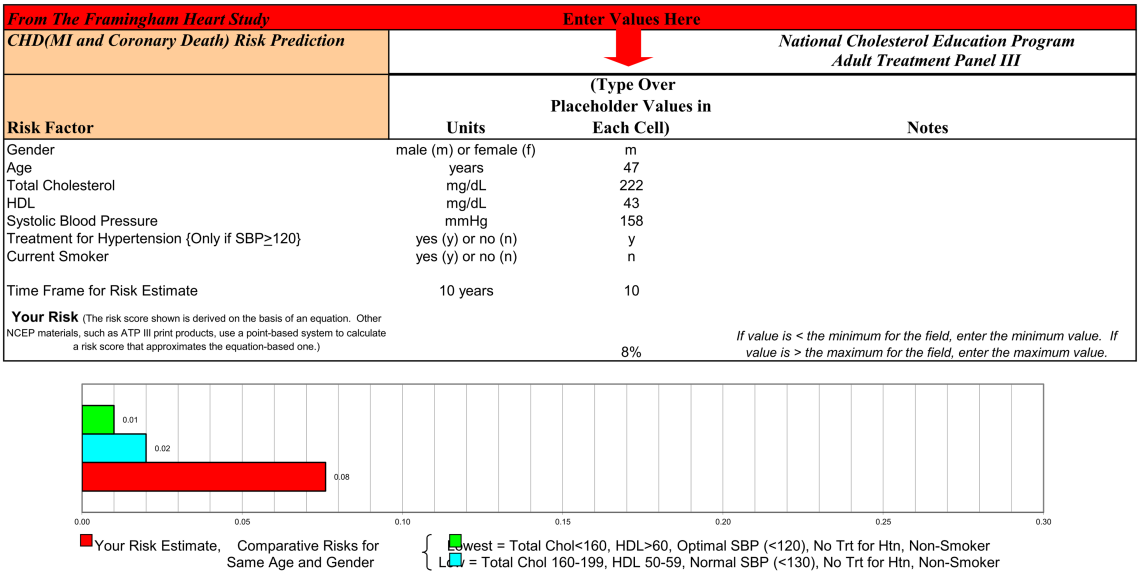

Figure D.1: FRS interface.

D.2 PROspective Cardiovascular Münster (PROCAM)

The PROCAM Quick Check estimates the 10-year risk of developing a coronary event. This version allows rapid initial assessment of coronary risk and includes only variables that do not require laboratory testing. The other version is a more detailed score that includes laboratory values (LDL-cholesterol, HDL-cholesterol, triglycerides and blood glucose). This platform is suitable for subjects aged 20-75 years. The interactive PROCAM Quick Check version is presented in Figure D.2.

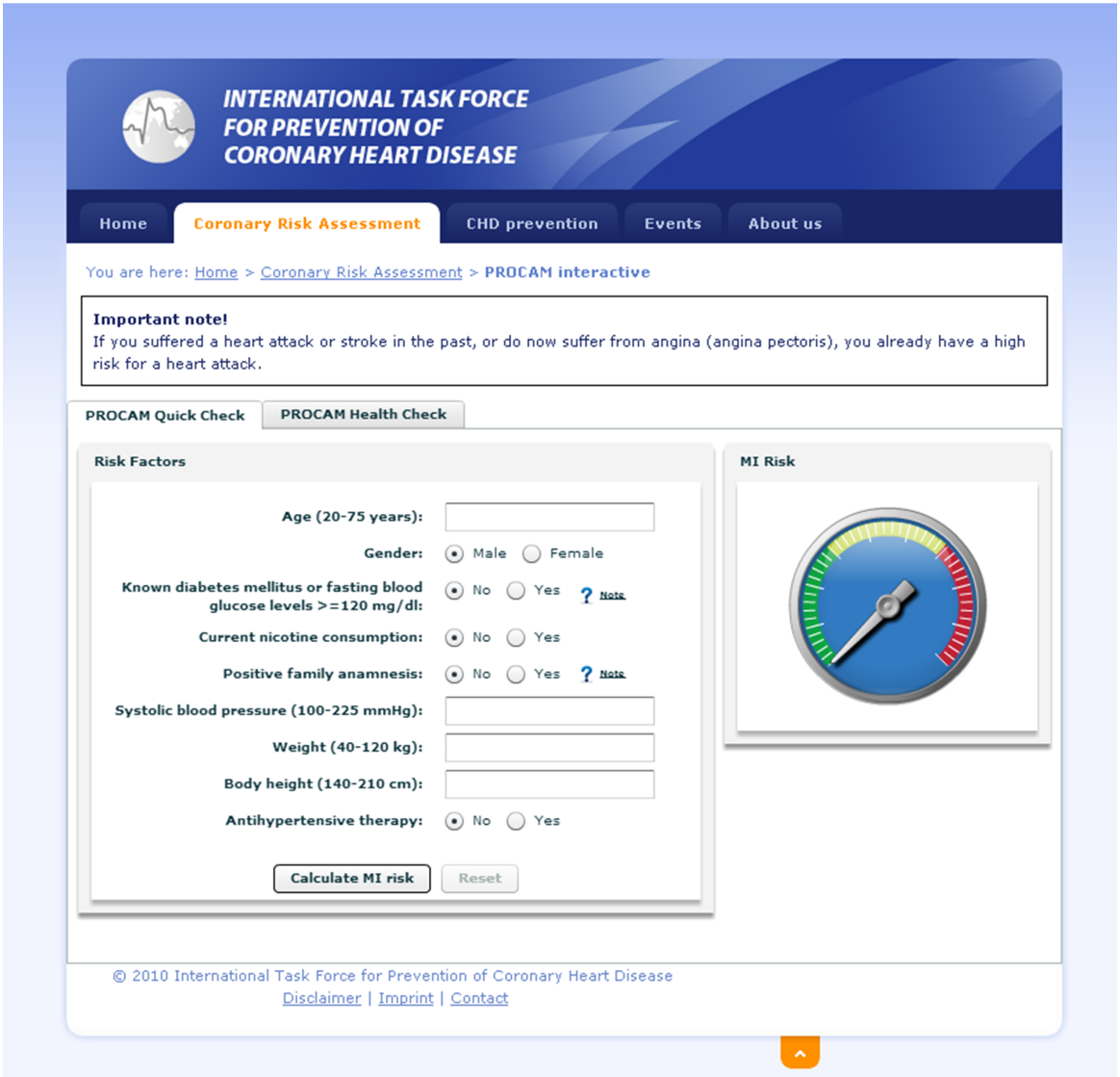

Figure D.2: PROCAM interactive tool.

D.3 ASsessing cardiovascular risk using Scottish Intercollegiate Guidelines Network (ASSIGN)

ASSIGN includes social deprivation for the first time, and family history of cardiovascular disease, with the classic risk factors. It identifies people free of cardiovascular disease most likely to develop it over 10-years. ”High risk” (score 20 or more) implies risk-lowering medication and/or other medical help. The ASSIGN interactive tool is presented in Figure D.3.

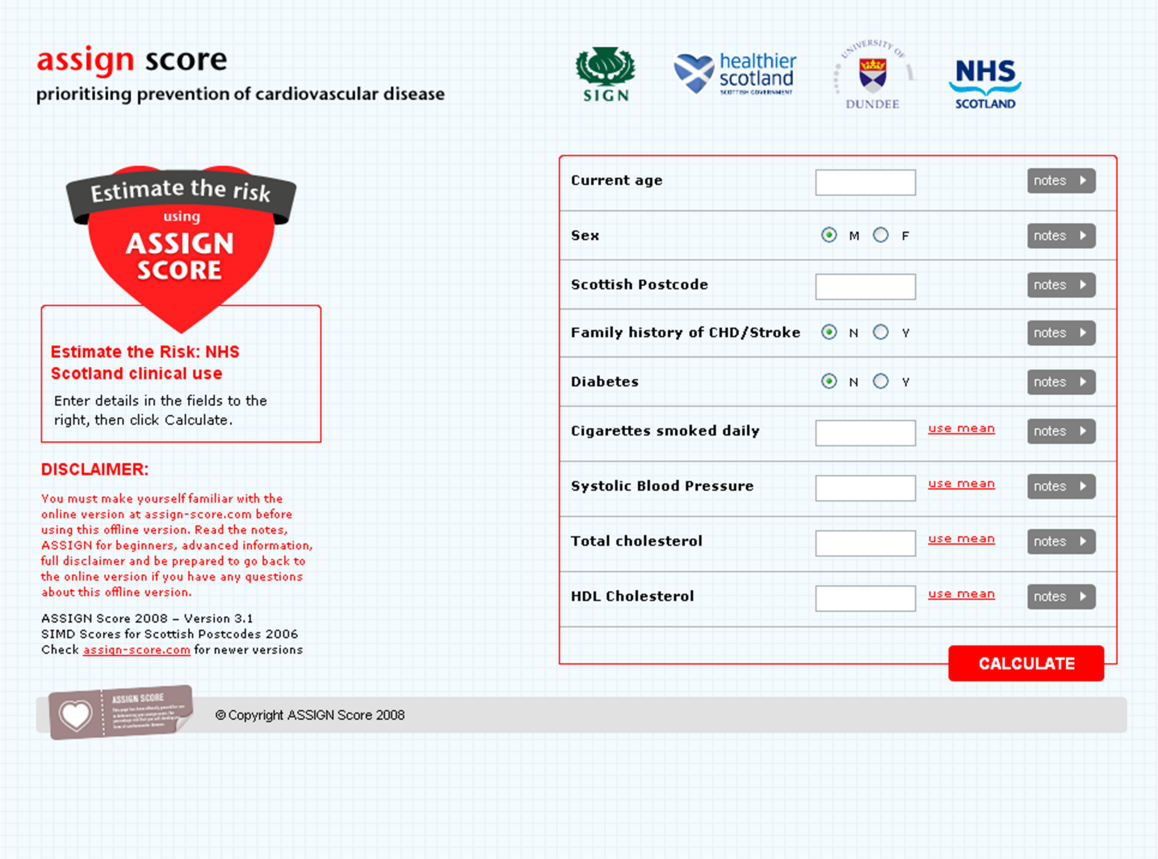

Figure D.3: ASSIGN interactive tool.

D.4 Systematic Coronary Risk Evaluation Project (SCORE)

SCORE is based on a large dataset tested thoroughly on European data. The SCORE risk function can be calibrated to each country's national mortality statistics. Two versions are available. The low risk charts can be used in Andorra, Austria, Belgium, Cyprus, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Israel, Italy, Luxembourg, Malta, Monaco, The Netherlands, Norway, Portugal, San Marino, Slovenia, Spain, Sweden, Switzerland and the United Kingdom. The high risk charts may be used other European countries. These include Armenia, Azerbaijan, Belarus, Bulgaria, Georgia, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Macedonia FYR, Moldova, Russia, Ukraine and Uzbekistan. CVD risk assessment tool is presented in Figure D.4.

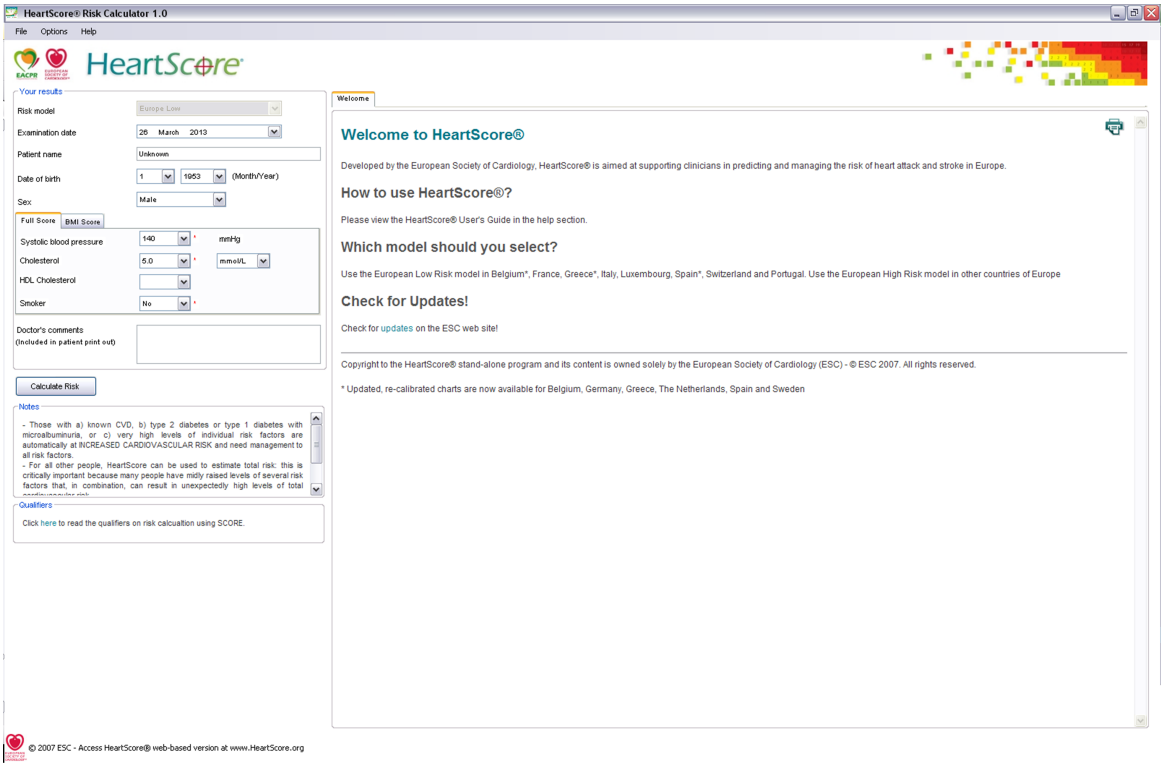

Figure D.4: SCORE interactive tool.

E Additional Statistical Data

E.1 Operators Significance

To complete the operator analysis, the categorization for operators and sub-categorization for carotid site was also performed and, it is presented in Table E.1.

Table E.1: Descriptive data categorized by operator and sub-categorized by carotid site.

Parameter	**Operator1**			**Operator2**			**Operator3**		
	Left	*Right*	*p*	*Left*	*Right*	*p*	*Left*	*Right*	*p*
SWTT (ms)	170.46±60.99	180.27±60.21	$< .01$	174.89±66.05	186.35±65.30	$< .01$	173.02±60.65	197.94±59.18	$< .01$
RWTT (ms)	162.88±44.08	150.36±43.10	$< .01$	161.16±35.81	151.45±44.70	$< .01$	162.73±41.49	154.38±44.11	$< .01$
LVET (ms)	278.17±40.61	284.25±39.09	$< .01$	288.11±41.64	289.12±46.66	0.11	285.54±39.15	294.87±36.44	$< .01$
RWA (a.u.[a])	0.86±0.12	0.86±0.12	0.06	0.85±0.10	0.85±0.11	$< .05$	0.85±0.09	0.84±0.12	0.51
DWA (a.u.[a])	0.76±0.12	0.77±0.11	0.64	0.75±0.12	0.79±0.12	$< .01$	0.72±0.14	0.79±0.12	$< .01$
AIx (%)	−0.10±18.6	5.03±17.44	$< .01$	−0.10±17.80	5.06±17.94	$< .01$	−1.3±17.86	7.19±19.18	$< .01$

[a] Arbitrary Amplitude Units

E.2 Sub-categorization by Subject

In this section are presented the GroupR (Trial2) results, considering each subject and operator individually, for the SWTT, RWTT, LVET, RWA, DWA, measured at the left and right carotid arteries.

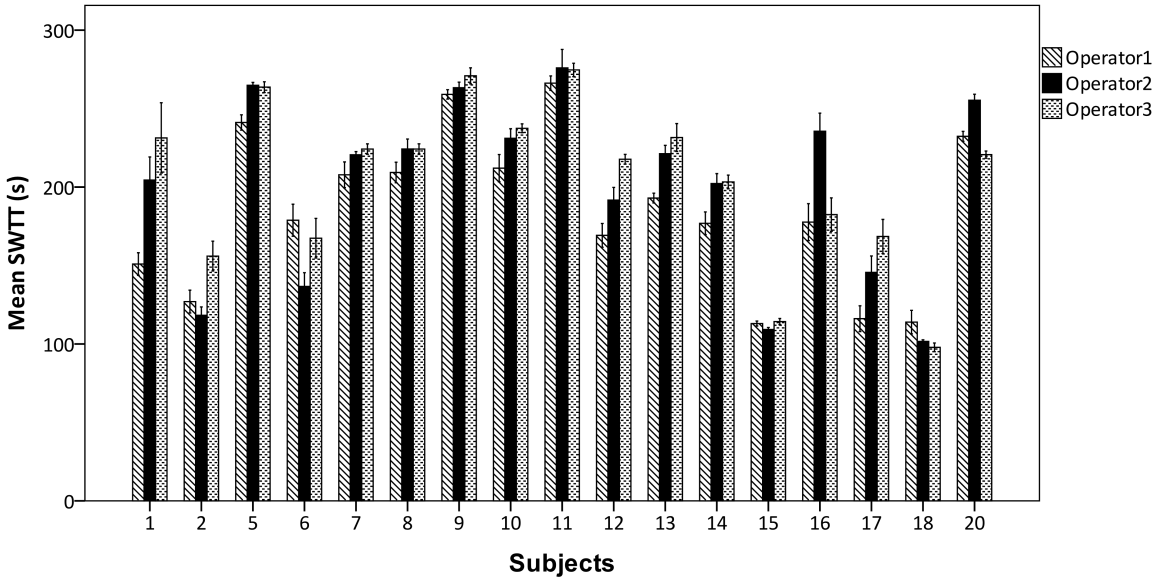

Figure E.1: Bar plot for the SWTT measured at right carotid site.

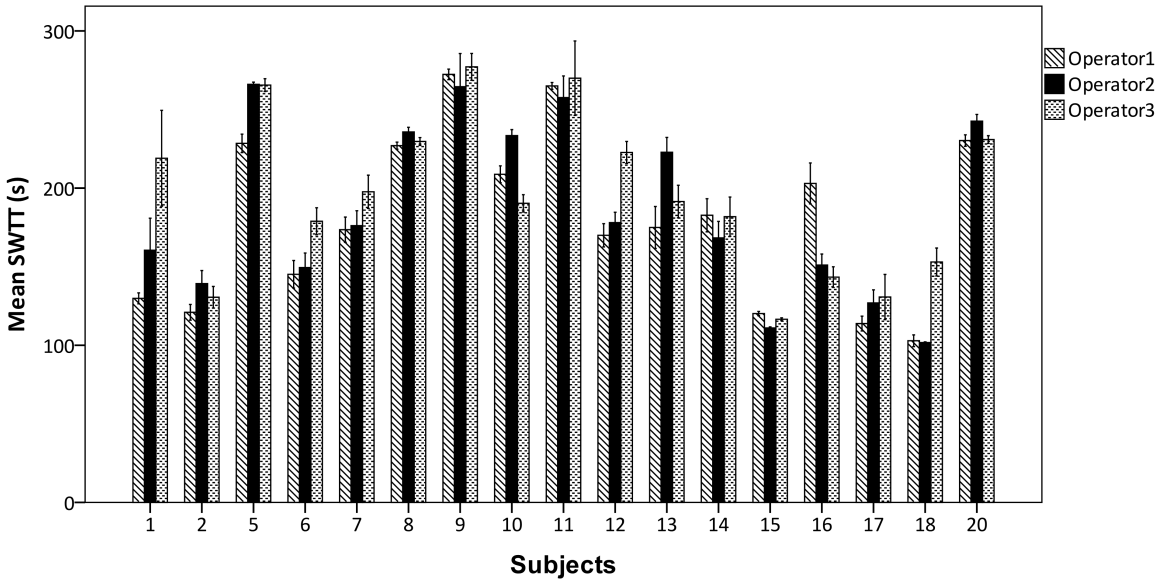

Figure E.2: Bar plot for the SWTT measured at left carotid site.

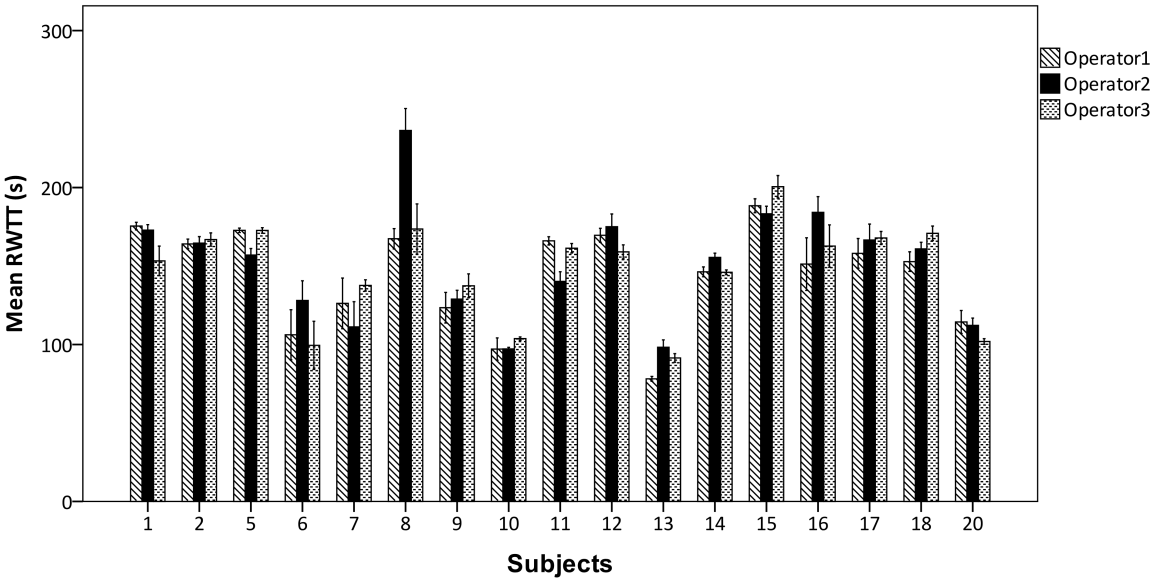

Figure E.3: Bar plot for the RWTT measured at right carotid site.

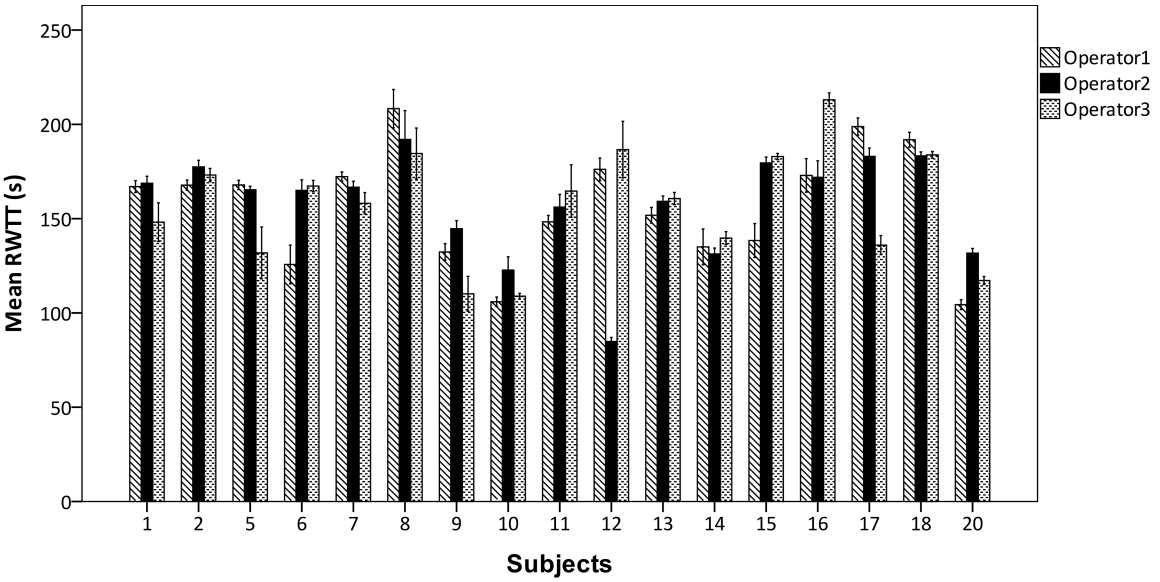

Figure E.4: Bar plot for the RWTT measured at left carotid site.

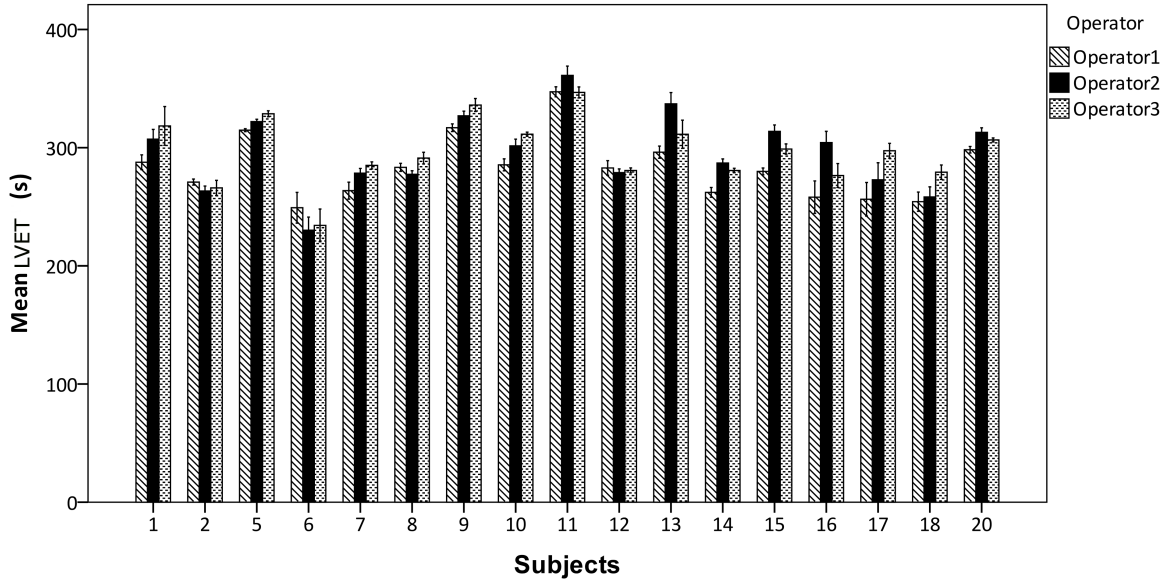

Figure E.5: Bar plot for the LVET measured at right carotid site.

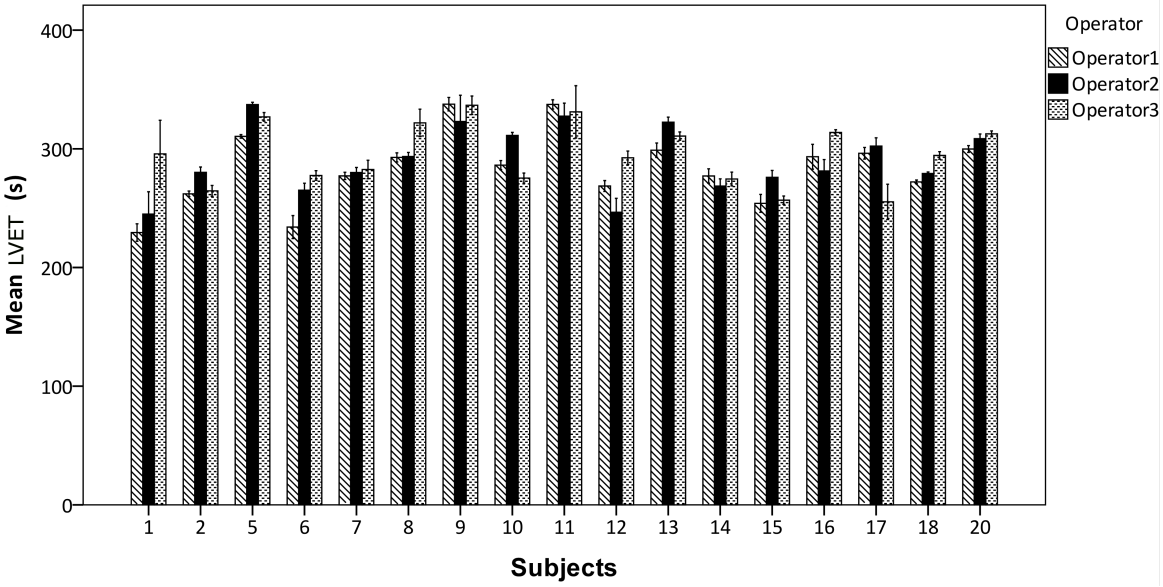

Figure E.6: Bar plot for the LVET measured at left carotid site.

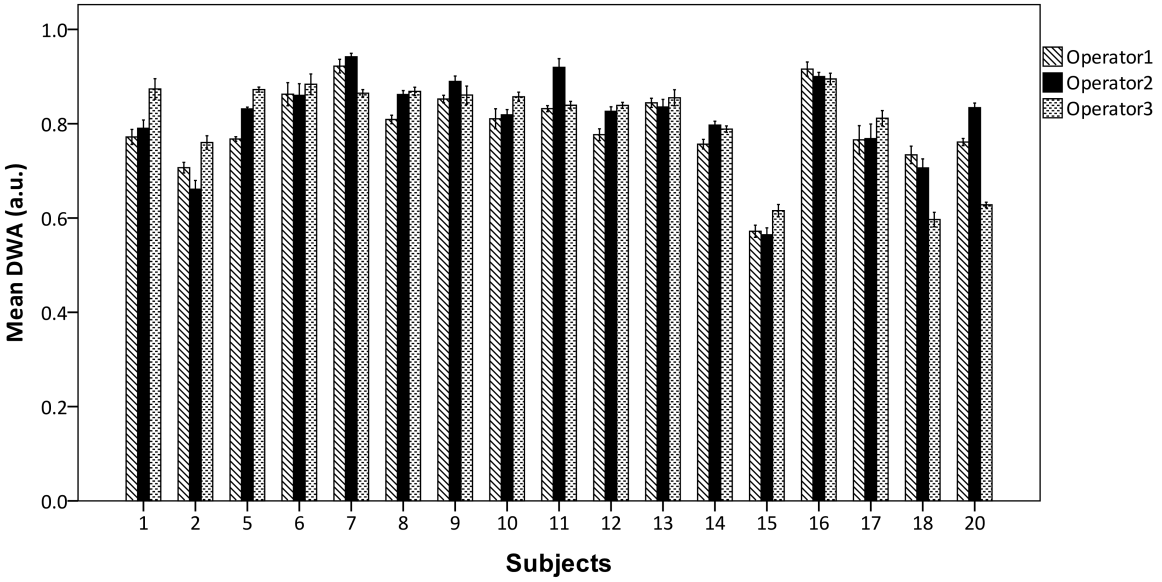

Figure E.7: Bar plot for the DWA measured at right carotid site.

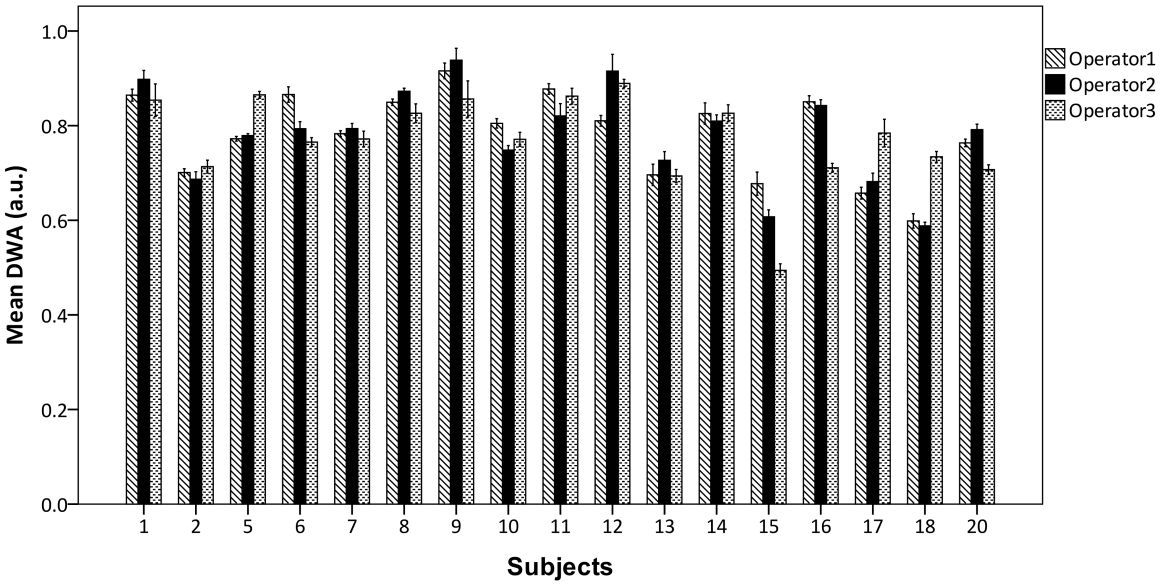

Figure E.8: Bar plot for the DWA measured at left carotid site.

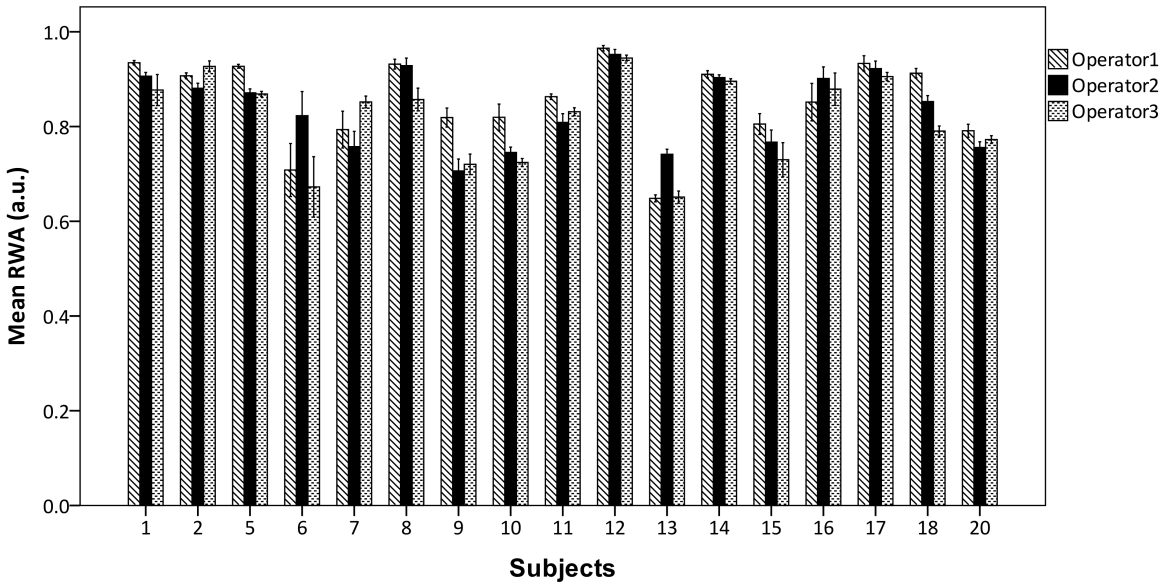

Figure E.9: Bar plot for the RWA measured at right carotid site.

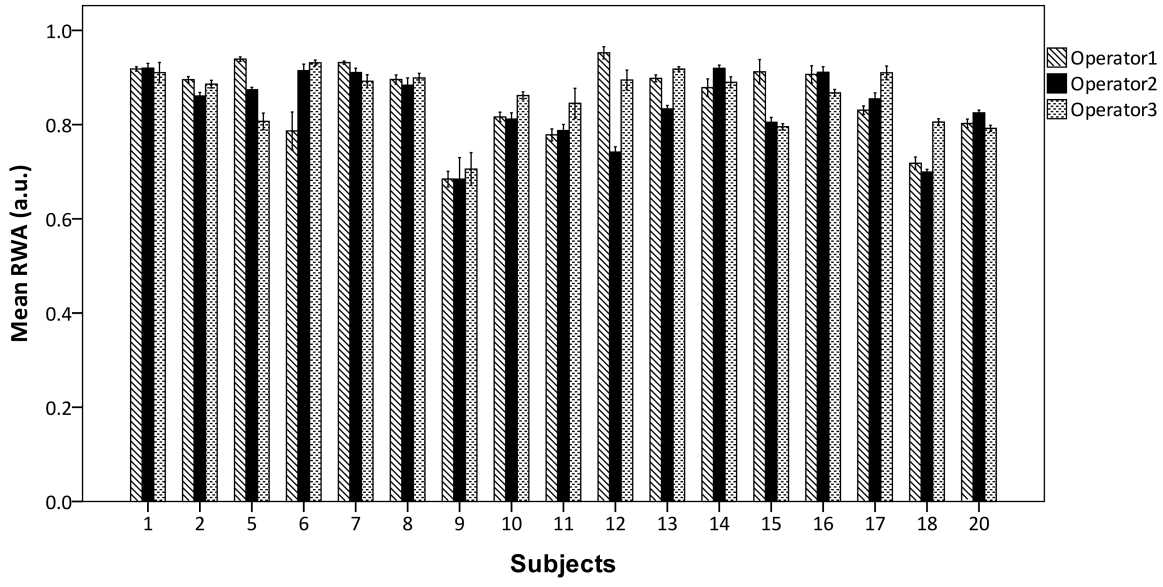

Figure E.10: Bar plot for the RWA measured at left carotid site.

E.3 Trials Significance

Table E.2: Comparison of variability values obtained from commercial devices.

Subject	Operator	Carotid	SWTT	RWTT	RWA	LVET	DWA	AIx
1	1	Left	0.16	$p < 0.01$	$p < 0.05$	0.66	0.06	$p < 0.01$
		Right	0.95	$p < 0.01$	$p < 0.05$	0.07	$p < 0.01$	$p < 0.01$
	2	Left	$p < 0.01$	$p < 0.05$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.05$
		Right	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$
	3	Left	$p < 0.05$	$p < 0.05$	$p < 0.05$	0.19	0.62	$p < 0.05$
		Right	0.21	0.34	0.71	0.10	0.41	0.34
2	1	Left	$p < 0.01$	$p < 0.01$	0.48	0.20	0.84	0.64
		Right	$p < 0.01$	$p < 0.01$	0.83	0.08	0.34	0.89
	2	Left	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.05$	0.07
		Right	0.18	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$
	3	Left	$p < 0.05$	$p < 0.01$	$p < 0.01$	$p < 0.05$	$p < 0.01$	$p < 0.01$
		Right	0.18	0.46	0.30	0.09	0.36	0.23
5	1	Left	$p < 0.05$	0.88	$p < 0.05$	$p < 0.01$	$p < 0.01$	0.08
		Right	$p < 0.01$	$p < 0.01$	0.52	$p < 0.01$	$p < 0.01$	0.68
	2	Left	0.10	0.37	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$
		Right	0.87	$p < 0.05$	$p < 0.05$	0.20	$p < 0.01$	$p < 0.05$
	3	Left	0.08	0.72	0.68	0.28	0.20	0.77
		Right	$p < 0.05$	0.86	$p < 0.01$	$p < 0.05$	$p < 0.01$	$p < 0.01$
6	1	Left	0.15	$p < 0.01$	$p < 0.01$	$p < 0.05$	$p < 0.05$	$p < 0.01$
		Right	$p < 0.05$	0.24	0.96	$p < 0.01$	$p < 0.05$	0.60
	2	Left	$p < 0.01$	$p < 0.01$	$p < 0.05$	0.30	$p < 0.01$	$p < 0.01$
		Right	0.21	$p < 0.01$	$p < 0.05$	$p < 0.05$	$p < 0.05$	$p < 0.01$
	3	Left	$p < 0.05$	$p < 0.01$	$p < 0.01$	$p < 0.05$	0.86	$p < 0.01$
		Right	$p < 0.01$	0.06	$p < 0.01$	$p < 0.01$	0.06	$p < 0.01$
7	1	Left	0.40	$p < 0.01$	$p < 0.01$	0.42	$p < 0.05$	$p < 0.01$
		Right	$p < 0.01$	0.36	0.87	$p < 0.01$	$p < 0.05$	0.82
	2	Left	0.32	0.06	0.07	$p < 0.01$	$p < 0.01$	0.85
		Right	$p < 0.05$	0.08	0.06	0.20	$p < 0.05$	0.06
	3	Left	$p < 0.01$	$p < 0.05$	0.13	$p < 0.01$	$p < 0.01$	$p < 0.05$
		Right	$p < 0.01$	0.75	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$
8	1	Left	0.14	$p < 0.05$	0.93	$p < 0.01$	$p < 0.05$	0.13
		Right	$p < 0.01$	0.44	0.19	$p < 0.01$	0.06	0.55
	2	Left	0.16	0.60	0.06	0.22	$p < 0.05$	0.66
		Right	0.65	0.62	0.72	0.64	0.47	0.42
	3	Left	0.47	$p < 0.05$	$p < 0.05$	$p < 0.05$	$p < 0.01$	0.06
		Right	0.14	0.58	0.58	0.46	0.77	0.91
9	1	Left	$p < 0.05$	$p < 0.05$	0.95	0.42	0.30	0.95
		Right	0.51	$p < 0.05$	0.05	0.40	0.45	$p < 0.05$
	2	Left	0.17	0.09	0.09	$p < 0.05$	0.07	0.09
		Right	$p < 0.05$	$p < 0.01$	$p < 0.01$	$p < 0.05$	$p < 0.01$	$p < 0.01$
	3	Left	0.12	0.27	0.33	0.24	0.69	0.33
		Right	$p < 0.01$	$p < 0.01$	$p < 0.01$	0.30	$p < 0.01$	$p < 0.01$
10	1	Left	$p < 0.05$	$p < 0.01$	$p < 0.05$	$p < 0.01$	0.35	$p < 0.05$
		Right	$p < 0.01$	0.07	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$
	2	Left	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$
		Right	$p < 0.01$	$p < 0.05$	0.55	$p < 0.01$	$p < 0.01$	0.54

Continued on next page

Table E.2 – *Continued from previous page*

Subject	Operator	Carotid	SWTT	RWTT	RWA	LVET	DWA	AIx
	3	Left	0.20	$p < 0.01$	$p < 0.01$	0.36	$p < 0.01$	$p < 0.01$
		Right	$p < 0.01$	$p < 0.01$	0.21	0.25	$p < 0.01$	0.21
11	1	Left	0.47	0.36	0.81	0.14	0.70	0.81
		Right	0.51	$p < 0.01$	0.40	$p < 0.01$	0.56	0.35
	2	Left	0.32	$p < 0.01$	$p < 0.01$	0.09	$p < 0.01$	$p < 0.01$
		Right	0.47	$p < 0.05$	0.63	0.53	$p < 0.05$	0.63
	3	Left	0.42	0.95	0.14	0.67	0.39	0.14
		Right	0.41	$p < 0.01$	$p < 0.01$	0.44	0.95	$p < 0.01$
12	1	Left	$p < 0.01$	$p < 0.05$	0.28	0.30	$p < 0.01$	$p < 0.01$
		Right	$p < 0.01$	$p < 0.01$	0.71	$p < 0.01$	$p < 0.05$	$p < 0.01$
	2	Left	0.52	0.73	0.35	0.37	0.29	0.35
		Right	$p < 0.01$	$p < 0.01$	$p < 0.01$	0.72	$p < 0.01$	$p < 0.01$
	3	Left	0.80	$p < 0.01$	0.51	$p < 0.01$	$p < 0.05$	$p < 0.01$
		Right	0.07	0.12	$p < 0.05$	0.83	$p < 0.05$	$p < 0.01$
13	1	Left	$p < 0.05$	$p < 0.05$	0.18	$p < 0.01$	$p < 0.01$	$p < 0.01$
		Right	0.89	0.40	0.30	0.40	$p < 0.01$	0.30
	2	Left	0.79	$p < 0.05$	$p < 0.01$	0.81	0.29	$p < 0.01$
		Right	$p < 0.01$	0.30	$p < 0.01$	$p < 0.05$	0.08	$p < 0.01$
	3	Left	0.15	$p < 0.01$	0.54	$p < 0.01$	$p < 0.01$	$p < 0.05$
		Right	$p < 0.01$	0.33	$p < 0.01$	$p < 0.01$	0.34	$p < 0.01$
14	1	Left	$p < 0.05$	$p < 0.01$	$p < 0.01$	0.21	$p < 0.01$	$p < 0.01$
		Right	0.73	0.35	0.75	0.53	0.06	0.51
	2	Left	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.05$	0.10	$p < 0.01$
		Right	$p < 0.01$	$p < 0.01$	0.08	$p < 0.01$	$p < 0.01$	$p < 0.01$
	3	Left	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	0.41	$p < 0.01$
		Right	0.81	0.23	0.36	0.85	$p < 0.05$	0.43
15	1	Left	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	0.43	$p < 0.01$
		Right	$p < 0.01$	0.20	$p < 0.05$	$p < 0.01$	$p < 0.05$	$p < 0.05$
	2	Left	0.16	$p < 0.05$	0.19	0.59	0.88	0.28
		Right	0.31	0.49	0.70	$p < 0.05$	0.17	0.70
	3	Left	0.08	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$
		Right	0.74	$p < 0.01$	$p < 0.01$	$p < 0.01$	0.40	$p < 0.01$
16	1	Left	0.12	0.22	$p < 0.05$	$p < 0.05$	0.24	0.85
		Right	$p < 0.05$	0.20	$p < 0.05$	0.45	0.41	0.29
	2	Left	0.55	0.07	0.78	0.09	0.06	$p < 0.05$
		Right	0.12	$p < 0.01$	$p < 0.05$	0.06	$p < 0.01$	$p < 0.05$
	3	Left	$p < 0.05$	$p < 0.01$	$p < 0.01$	$p < 0.01$	0.77	$p < 0.05$
		Right						
17	1	Left	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$
		Right	0.08	0.20	0.61	0.30	0.14	0.38
	2	Left	0.49	$p < 0.05$	$p < 0.01$	0.77	$p < 0.01$	0.07
		Right	0.13	$p < 0.05$	0.12	$p < 0.05$	$p < 0.05$	$p < 0.01$
	3	Left	0.08	0.11	0.11	0.84	0.19	0.94
		Right	0.97	0.52	0.28	0.13	0.62	0.52
18	1	Left	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$
		Right	$p < 0.01$	$p < 0.01$	$p < 0.05$	$p < 0.01$	$p < 0.01$	$p < 0.01$
	2	Left	$p < 0.01$	0.79	0.07	$p < 0.05$	$p < 0.05$	$p < 0.05$
		Right	$p < 0.01$	$p < 0.01$	0.71	$p < 0.01$	$p < 0.01$	0.71
	3	Left	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$
		Right	$p < 0.01$	$p < 0.01$	$p < 0.05$	$p < 0.01$	$p < 0.01$	$p < 0.01$

Continued on next page

Table E.2 – *Continued from previous page*

Subject	Operator	Carotid	SWTT	RWTT	RWA	LVET	DWA	AIx
20	1	Left	0.16	$p < 0.01$	$p < 0.01$	0.38	$p < 0.01$	$p < 0.01$
		Right	$p < 0.01$	$p < 0.01$	0.50	$p < 0.05$	$p < 0.01$	0.25
	2	Left	0.15	$p < 0.01$	$p < 0.01$	$p < 0.05$	$p < 0.05$	$p < 0.01$
		Right	0.44	0.36	0.13	0.27	0.22	0.13
	3	Left	0.30	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$
		Right	$p < 0.01$	$p < 0.01$	$p < 0.01$	0.43	$p < 0.01$	$p < 0.01$

References

[1] P. Boutouyrie, A. I. Tropeano, R. Asmar, I. Gautier, A. Benetos, P. Lacolley, and L. Laurent. Aortic stiffness is an independent predictor of primary coronary events in hypertensive patients: a longitudinal study. *Hypertension*, 39(1):10–15, 2002. 2, 21

[2] S. Laurent et al. Expert consensus document on arterial stiffness methodological issues and clinical applications. *European Heart Journal*, 27(21):2588–2605, 2006. 2, 10, 11, 81

[3] P. K. Hamilton, C. J. Lockhart, C. E. Quinn, and G. E. McVeigh. Arterial stiffness: clinical relevance, measurement and treatment. *Clin Sci (Lond)*, 113(4):157–70, 2007. 2

[4] S. Laurent and P. Boutouyrie. Arterial stiffness and cardiovascular events in hypertensives. *Current Cardiovascular Risk Reports*, 7(3):238–243, 2013. 2, 11

[5] P. Boutouyrie et al. Determinants of pulse wave velocity in healthy people and in the presence of cardiovascular risk factors: establishing normal and reference values. *European Heart Journal*, 31:2338–2350, 2010. 2

[6] D. A. Duprez and J. N. Cohn. Arterial stiffness as a risk factor for coronary atherosclerosis. *Coronary Heart Disease*, 9:139144, 2007. 2, 20

[7] E. Kis et al. Pulse wave velocity in end-stage renal disease: influence of age and body dimensions. *Pediatric Research*, 63, 2008. 2

[8] K. Cruickshank, L. Riste, S. G. Anderson, J. S. Wright, G. Dunn, and R. G. Gosling. Aortic pulse-wave velocity and its relationship to mortality in diabetes and glucose intolerance: an integrated index of vascular function? *Circulation*, 106(16):2085–2090, 2002. 2

[9] A. P. Avolio, M. Butlin, and A. Walsh. Arterial blood pressure measurement and pulse wave analysis–their role in enhancing cardiovascular assessment. *Physiol Meas*, 31(1):R1–47, 2010. 2, 18, 19, 21, 27, 81

[10] A. L. Pauca, M. F. O'Rourke, and N. D. Kon. Prospective evaluation of a method for estimating ascending aortic pressure from the radial artery pressure waveform. *Hypertension*, 38(4):932–937, 2001. 2

[11] V. Kotsis, S. Stabouli, I. Karafillis, S. Papakatsika, Z. Rizos, S. Miyakis, S. Goulopoulou, G. Parati, and P. Nilsson. Arterial stiffness and 24 h ambulatory blood pressure monitoring in young healthy volunteers: the early vascular ageing Aristotle University Thessaloniki Study (EVA-ARIS Study). *Atherosclerosis*, 219(1):194–9, 2011. 2, 18

[12] K. H. Parker. A brief history of arterial wave mechanics. *Med Biol Eng Comput*, 47(2):111–8, 2009. 2

[13] L. M. Van Bortel et al. Expert consensus document on the measurement of aortic stiffness in daily practice using carotid-femoral pulse wave velocity. *J Hypertens*, 30(3):445–8, 2012. 2, 21

[14] J. Nurnberger, A. Keflioglu-Scheiber, A. M. Opazo Saez, R. R. Wenzel, T. Philipp, and R. F. Schafers. Augmentation index is associated with cardiovascular risk. *J Hypertens*, 20(12):2407–14, 2002. 2

[15] I. Yoo, P. Alafaireet, M. Marinov, K. Pena-Hernandez, R. Gopidi, J. Chang, and L. Hua. Data mining in healthcare and biomedicine: a survey of the literature. *J. Med. Syst*, 2011. 3, 104, 110

[16] M. Kumari and S. Godara. Comparative study of data mining classification methods in cardiovascular disease prediction. *IJCST*, 2(2):303–308, 2011. 3

[17] S. Paredes, T. Rocha, P. de Carvalho, J. Henriques, M. Harris, and J. Morais. Long term cardiovascular risk models' combination. *Comput Methods Programs Biomed*, 101(3):231–42, 2011. 3, 35

[18] B. Pannier, A. Avolio, A. Hoeks, G. Mancia, and k. Takazawa. Methods and devices for measuring arterial compliance in humans. *AJH*, 15:743–753, 2002. 3, 20

[19] S. Laurent et al. Expert consensus document on arterial stiffness: methodological issues and clinical applications. *European Heart Journal*, 27:25882605, 2006. 3, 10

[20] A. Swillens and P. Segers. Assessment of arterial pressure wave reflection: methodological considerations. *Artery Research*, 2(4):122–131, 2008. 3

[21] S. Mendis, P. Puska, and B. Norrving. *Global atlas on cardiovascular disease prevention and control*. World Health Organization, Geneva, 2011. 10, 12

[22] T. Weber, J. Auer, M. F. O'Rourke, E. Kvas, E. Lassnig, R. Berent, and B. Eber. Arterial stiffness, wave reflections, and the risk of coronary artery disease. *Circulation*, 109(2):184–9, 2004. 10

[23] S. J. Zieman, V. Melenovsky, and D. A. Kass. Mechanisms, pathophysiology, and therapy of arterial stiffness. *Arterioscler Thromb Vasc Biol*, 25(5):932–43, 2005. 10

[24] G. M. London and A. P. Guerin. Influence of arterial pulse and reflected waves on blood pressure and cardiac function. *Am Heart J*, 138(3):220–223, 1999. 10

[25] J. D. Cameron, B. P. McGrath, and A. M. Dart. Use of radial artery applanation tonometry and a generalized transfer function to determine aortic pressure augmentation in subjects with treated hypertension. *JACC*, 32(5):1214–20, 1998. 10

[26] I. Ikonomidis, S. Tzortzis, T. Papaioannou, A. Protogerou, K. Stamatelopoulos, C. Papamichael, N. Zakopoulos, and J. Lekakis. Incremental value of arterial wave reflections in the determination of left ventricular diastolic dysfunction in untreated patients with essential hypertension. *J Hum Hypertens*, 22(10):687–98, 2008. 10

[27] M. Gordon. Heart and vessels. http://www.studyblue.com/notes/note/n/heart-and-vessels/deck/6409908, 2013. 11

[28] I. S. Mackenzie, I. B. Wilkinson, and J. R. Cockroft. Assessment of arterial stiffness in clinical practice. *Q J Med.*, 95:67–74, 2002. 11

[29] M. T. Cooney, H. C. Cooney, A. Dudina, and I. M. Graham. Assessment of cardiovascular risk. *Curr Hypertens Rep*, 12(5):384–93, 2010. 11

[30] W. W. Nichols and M. F. O'Rourke. *McDonald's blood flow in arteries: theoretical, experimental and clinical principles.* Hodden Arnold, London, fifth edition, 2005. 11, 22

[31] M. Juonala et al. Life-time risk factors and progression of carotid atherosclerosis in young adults: the Cardiovascular Risk in Young Finns study. *Eur Heart J*, 31(14):1745–51, 2010. 12

[32] D. Lloyd-Jones et al. Heart disease and stroke statistics–2010 update: a report from the american heart association. *Circulation*, 121(7):e46–e215, 2010. 12

[33] J. Dernellis and M. Panaretou. Aortic stiffness is an independent predictor of progression to hypertension in nonhypertensive subjects. *Hypertension*, 45(3):426–31, 2005. 12

[34] D. Liao, D. K. Arnett, H. A. Tyroler, W. A. Riley, L. E. Chambless, M. Szklo, and G. Heiss. Arterial stiffness and the development of hypertension : the ARIC study. *Hypertension*, 34(2):201–206, 1999. 12

[35] R. A. Payne, I. B. Wilkinson, and D. J. Webb. Arterial stiffness and hypertension: emerging concepts. *Hypertension*, 55(1):9–14, 2010. 12

[36] S. S. Najjar, A. Scuteri, and E. G. Lakatta. Arterial aging: is it an immutable cardiovascular risk factor? *Hypertension*, 46(3):454–62, 2005. 13

[37] M. F. O'Rourke and J. Hashimoto. Mechanical factors in arterial aging: a clinical perspective. *J Am Coll Cardiol*, 50(1):1–13, 2007. 13

[38] M. F. O'Rourke and W. W. Nichols. Aortic diameter, aortic stiffness, and wave reflection increase with age and isolated systolic hypertension. *Hypertension*, 45(4):652–8, 2005. 13

[39] H. Y. Lee and B. H. Oh. Aging and arterial stiffness. *Circulation Journal*, 74(11):2257–2262, 2010. 13

[40] J. Perk et al. European guidelines on cardiovascular disease prevention in clinical practice (version 2012). *Eur Heart J*, 33(13):1635–701, 2012. 13, 14, 15, 16, 20, 30, 31, 34, 100

[41] T. A. Mulders, B. van den Bogaard, A. Bakker, M. D. Trip, E. S. Stroes, B. J. van den Born, and S. J. Pinto-Sietsma. Arterial stiffness is increased in families with premature coronary artery disease. *Heart*, 98(6):490–4, 2012. 13

[42] M. A. Albert, R. J. Glynn, J. Buring, and P. M. Ridker. Impact of traditional and novel risk factors on the relationship between socioeconomic status and incident cardiovascular events. *Circulation*, 114(24):2619–26, 2006. 13

[43] S. Stringhini, S. Sabia, M. Shipley, E. Brunner, H. Nabi, M. Kivimaki, and A. Singh-Manoux. Association of socioeconomic position with health behaviors and mortality. *JAMA*, 303(12):1159–66, 2010. 13

[44] M. C. Carr and J. D. Brunzell. Abdominal obesity and dyslipidemia in the metabolic syndrome: importance of Type 2 diabetes and familial combined hyperlipidemia in coronary artery disease risk. *J Clin Endocrinol Metab*, 89(6):2601–7, 2004. 14

[45] B. Hacihamdioglu, V. Okutan, Y. Yozgat, D. Yildirim, M. Kocaoglu, M. K. Lenk, and O. Ozcan. Abdominal obesity is an independent risk factor for increased carotid intima-media thickness in obese children. *Turk J Pediatr.*, 53(1):4854, 2011. 14

[46] A. Nemes, H. Gavaller, E. Csajbok, T. Forster, and M. Csanady. Obesity is associated with aortic enlargement and increased stiffness: an echocardiographic study. *Int J Cardiovasc Imaging*, 24(2):165–71, 2008. 14

[47] S. E. Kahn, B. Zinman, S. M. Haffner, M. C. O'Neill, B. G. Kravitz, D. Yu, M. I. Freed, W. H. Herman, R. R. Holman, N. P. Jones, J. M. Lachin, and G. C. Viberti. Obesity is a major determinant of the association of C-reactive protein levels and the metabolic syndrome in type 2 diabetes. *Diabetes*, 55(8):2357–64, 2006. 14

[48] M. E. Safar, B. I. Levy, and H. Struijker-Boudier. Current perspectives on arterial stiffness and pulse pressure in hypertension and cardiovascular diseases. *Circulation*, 107(22):2864–9, 2003. 14

[49] R. S. Vasan, M. G. Larson, E. P. Leip, W. B. Kannel, and D. Levy. Assessment of frequency of progression to hypertension in non-hypertensive participants in the Framingham Heart Study: a cohort study. *The Lancet*, 358(9294):1682–1686, 2001. 14

[50] . S. Franklin and N. D. Wong. Hypertension and cardiovascular disease: contributions of the Framingham Heart Study. *Global Heart*, 8(1):49–57, 2013. 14

[51] I. Graham et al. European guidelines on cardiovascular disease prevention in clinical practice: executive summary. *Eur Heart J*, 28(19):2375–414, 2007. 14, 15

[52] M Laakso. Hyperglycemia and cardiovascular disease in Type 2 diabetes. *Diabetes*, 48(5):937–942, 1999. 14

[53] J. Kuusisto and M. Laakso. Update on Type 2 diabetes as a cardiovascular disease risk equivalent. *Curr Cardiol Rep*, 15(2):331, 2013. 14

[54] L. Ryden et al. Guidelines on diabetes, pre-diabetes, and cardiovascular diseases: executive summary. *Eur Heart J*, 28(1):88–136, 2007. 15

[55] M. Huisman, A. E. Kunst, and J. P. Mackenbach. Inequalities in the prevalence of smoking in the european union: comparing education and income. *Prev Med*, 40(6):756–64, 2005. 15

[56] Richard Edwards. The problem of tobacco smoking. *BMJ*, 328(7433):217–219, 2004. 15

[57] C. Baigent et al. Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90056 participants in 14 randomised trials of statins. *The Lancet*, 366(9493):1267–1278, 2005. 15

[58] Z. Reiner et al. ESC/EAS guidelines for the management of dyslipidaemias: the Task Force for the management of dyslipidaemias of the European Society of Cardiology (ESC) and the European Atherosclerosis Society (EAS). *Eur Heart J*, 32(14):1769–818, 2011. 16

[59] C. Baigent et al. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170000 participants in 26 randomised trials. *The Lancet*, 376(9753):1670 – 1681, 2010. 16

[60] M. T. Cooney, A. Dudina, D. De Bacquer, L. Wilhelmsen, S. Sans, A. Menotti, G. De Backer, P. Jousilahti, U. Keil, T. Thomsen, P. Whincup, and I. M. Graham. HDL cholesterol protects against cardiovascular disease in both genders, at all ages and at all levels of risk. *Atherosclerosis*, 206(2):611–6, 2009. 16

[61] M. J. Chapman et al. Triglyceride-rich lipoproteins and high-density lipoprotein cholesterol in patients at high risk of cardiovascular disease: evidence and guidance for management. *Eur Heart J*, 32(11):1345–61, 2011. 16

[62] J. E. Hokanson and M. A. Austin. Plasma triglyceride level is a risk factor for cardiovascular disease independent of high-density lipoprotein cholesterol level: a meta-analysis of population-based prospective studies. *Journal of Cardiovascular Risk*, 3(2):213–219, 1996. 16

[63] E. Di Angelantonio, N. Sarwar, P. Perry, S. Kaptoge, K. K. Ray, A. Thompson, A. M. Wood, S. Lewington, N. Sattar, C. J. Packard, R. Collins, S. G. Thompson, and J. Danesh. Major lipids, apolipoproteins, and risk of vascular disease. *JAMA*, 302(18):1993–2000, 2009. 16

[64] W. Lieb, M. G. Larson, E. J. Benjamin, X. Yin, G. H. Tofler, J. Selhub, P. F. Jacques, T. J. Wang, J. A. Vita, D. Levy, R. S. Vasan, and G. F. Mitchell. Multimarker approach to evaluate correlates of vascular stiffness: the Framingham Heart Study. *Circulation*, 119(1):37–43, 2009. 16, 17

[65] T. J. Wang, P. Gona, and M. G. Larson. Multiple biomarkers for the prediction of first major cardiovascular events and death. *N Engl J Med*, 355:2631–9, 2006. 16

[66] A. Thompson. Lipoprotein-associated phospholipase A(2) and risk of coronary disease, stroke, and mortality: collaborative analysis of 32 prospective studies. *The Lancet*, 375(9725):1536 – 1544, 2010. 17

[67] A. Mahmud and J. Feely. Arterial stiffness and the renin-angiotensin-aldosterone system. *J Renin Angiotensin Aldosterone Syst*, 5(3):102–8, 2004. 17

[68] C. M. Dias. 25 Anos de inquérito nacional de saúde em Portugal. *Revista Portuguesa de Saúde Pública*, pages 51–60, 2009. 17

[69] P. Perdigão, E. Rocha, J.S. Duarte, A. Santos, and A. Macedo. Prevalência, caracterização e distribuição dos principais factores de risco cardiovascular em Portugal. Uma análise do estudo AMÁLIA. *Rev Port Cardiol*, 30(4):393–432, 2011. 17

[70] E. O'Brien et al. European society of hypertension recommendations for conventional, ambulatory and home blood pressure measurement. *J Hypertens*, 21(5):821–48, 2003. 18

[71] H. C. Pereira. *Methods and instrumentation for cardiovascular condition assessment*. PhD thesis, University of Coimbra, 2013. 18

[72] D. A. Calhoun et al. Resistant hypertension: diagnosis, evaluation, and treatment. a scientific statement from the american heart association professional education committee of the council for high blood pressure research. *Hypertension*, 51(6):1403–19, 2008. 18

[73] J. Mesquita-Bastos, S. Bertoquini, and J. Polonia. Cardiovascular prognostic value of ambulatory blood pressure monitoring in a Portuguese hypertensive population followed up for 8.2 years. *Blood Press Monit*, 15(5):240–6, 2010. 18

[74] E. Dolan et al. Superiority of ambulatory over clinic blood pressure measurement in predicting mortality: the Dublin outcome study. *Hypertension*, 46(1):156–61, 2005. 18

[75] T. W. Hansen, J. Jeppesen, S. Rasmussen, H. Ibsen, and C. Torp-Pedersen. Ambulatory blood pressure and mortality: a population-based study. *Hypertension*, 45(4):499–504, 2005. 18

[76] P. Shaltis, A. Reisner, and H. Asada. Calibration of the photoplethysmogram to arterial blood pressure: capabilities and limitations for continuous pressure monitoring. In *Engineering in Medicine and Biology 27th Annual Conference*, pages 3970–3973, 2005. 18

[77] J. H. Song, J. S. Cho, H. S. Oh, J.S. Lee, and I. Y. Kim. Estimation of blood pressure using photoplethysmography on the wrist. *Computers in Cardiology*, 36:741–744, 2009. 18

[78] I. Jeong, S. Jun, D. Um, J. Oh, and H. Yoon. Non-invasive estimation of systolic blood pressure and diastolic blood pressure using photoplethysmograph components. *Yonsei Med J*, 51(3):345–53, 2010. 18

[79] J. S. Eckerle. *Encyclopedia of medical devices and instrumentation*, volume 6, chapter Arterial Tonometry, pages 402–8. John Wiley & Sons Inc, New York, 1988. 19

[80] J. Lee and K. C. Nam. *Biomedical Engineering*, chapter Tonometric Vascular Function Assessment, pages 549–66. InTech, 2009. 19

[81] J. Allen. Photoplethysmography and its application in clinical physiological measurement. *Physiol Meas*, 28(3):R1–39, 2007. 19

[82] P. M. Middleton, G. S. Chan, E. O'Lone, E. Steel, R. Carroll, B. G. Celler, and N. H. Lovell. Changes in left ventricular ejection time and pulse transit time derived from finger photoplethysmogram and electrocardiogram during moderate haemorrhage. *Clin Physiol Funct Imaging*, 29(3):163–9, 2009. 19

[83] S. R. Alty, N. Angarita-Jaimes, S. C. Millasseau, and P. J. Chowienczyk. Predicting arterial stiffness from the digital volume pulse waveform. *IEEE transactions on bio-medical engineering*, 54(12):2268–2275, 2007. 19

[84] T. Pereira, T. Oliveira, M. Cabeleira, H. Pereira, V. Almeida, J. Cardoso, and C. Correia. Comparison of low-cost and noninvasive optical sensors for cardiovascular monitoring. *Sensors Journal, IEEE*, 13(5):1434–1441, 2013. 19, 38

[85] S. Kohsaka and A. N. Makaryus. Coronary angiography using noninvasive imaging techniques of cardiac CT and MRI. *Current Cardiology Reviews*, 4:323–330, 2008. 20

[86] A. Harloff, T. Zech, F. Wegent, C. Strecker, C. Weiller, and M. Markl. Comparison of blood flow velocity quantification by 4D flow MR imaging with ultrasound at the carotid bifurcation. *AJNR Am J Neuroradiol*, 2013. 20

[87] X. Luo, Y. Yang, T. Cao, and Z. Li. Differences in left and right carotid intima-media thickness and the associated risk factors. *Clin Radiol*, 66(5):393–8, 2011. 20

[88] Arnold P. G. Hoeks, Christine Willekes, Pierre Boutouyrie, Peter J. Brands, Jean M. Willigers, and Robert S. Reneman. Automated detection of local artery wall thickness based on M-line signal processing. *Ultrasound in Medicine & Biology*, 23(7):1017–1023, 1997. 20

[89] J. M. Meinders and A. P. Hoeks. Simultaneous assessment of diameter and pressure waveforms in the carotid artery. *Ultrasound Med Biol*, 30(2):147–54, 2004. 20, 22

[90] S. J. Vermeersch, E. R. Rietzschel, M. L. De Buyzere, D. De Bacquer, G. De Backer, L. M. Van Bortel, T. C. Gillebert, P. R. Verdonck, and P. Segers. Determining carotid artery pressure from scaled diameter waveforms: comparison and validation of calibration techniques in 2026 subjects. *Physiol Meas*, 29(11):1267–1280, 2008. 20, 22

[91] Luc M. Van Bortel, Elisabeth J. Balkestein, Janneke J. van der Heijden-Spek, Floris H. Vanmolkot, Jan A. Staessen, Johannes A. Kragten, Jan W. Vredeveld, Michel E. Safar, Harry A. Struijker Boudier, and Arnold P. Hoeks. Non-invasive assessment of local arterial pulse pressure: comparison of applanation tonometry and echo-tracking. *J Hypertens*, 19(6):1037–1044, 2001. 20

[92] T. Pereira, T. Oliveira, M. Cabeleira, P. Matos, H. C. Pereira, VÀlmeida, E. Borges, H. Santos, T. Pereira, J. Cardoso, and C. Correia. Signal analysis in a new optical pulse waveform profiler for cardiovascular applications. In *Signal and Image Processing and Applications (SIPA 2011)*, pages 19–25, 2011. 20

[93] J. L. Cavalcante, J. A. Lima, A. Redheuil, and M. H. Al-Mallah. Aortic stiffness: current understanding and future directions. *J Am Coll Cardiol*, 57(14):1511–22, 2011. 20

[94] C. Vlachopoulos, K. Aznaouridis, and C. Stefanadis. Prediction of cardiovascular events and all-cause mortality with arterial stiffness: a systematic review and meta-analysis. *J Am Coll Cardiol*, 55(13):1318–27, 2010. 21

[95] S. Laurent, P. Boutouyrie, R. Asmar, I. Gautier, B. Laloux, L. Guize, P. Ducimetiere, and A. Benetos. Aortic stiffness is an independent predictor of all-cause and cardiovascular mortality in hypertensive patients. *Hypertension*, 37(5):1236–1241, 2001. 21

[96] H. C. Pereira, T. Pereira, V. Almeida, E. Borges, E. Figueiras, J. B. Simoes, J. L. Malaquias, J. M. Cardoso, and C. M. Correia. Characterization of a double probe for local pulse wave velocity assessment. *Physiol Meas*, 31(11):1449–65, 2010. 21, 38, 42, 69

[97] T. Pereira, M. Cabeleira, P. Matos, E. Borges, V. Almeida, Pereira H.C., Cardoso J.M., and Correia C. Non-contact pulse wave velocity assessment using optical methods. *Springer-Verlag Berlin Heidelberg*, pages 246–257, 2012. 21

[98] J. Kips, F. Vanmolkot, D. Mahieu, S. Vermeersch, I. Fabry, J. de Hoon, L. Van Bortel, and P. Segers. The use of diameter distension waveforms as an alternative for tonometric pressure to assess carotid blood pressure. *Physiol Meas*, 31(4):543–53, 2010. 22

[99] E. Hermeling, K. D. Reesink, R. S. Reneman, and A. P. G. Hoeks. Measurement of local pulse wave velocity: effects of signal processing on precision. *Ultrasound in Medicine & Biology*, 33(5):774–781, 2007. 22

[100] P. J. Brands, J. M. Willigers, L. A. F. Ledoux, R. Reneman, and A. P. G. Hoeks. A noninvasive method to estimate pulse wave velocity in arteries locally by means of ultrasound. *Ultrasound in Medicine & Biology*, 24(9):1325 – 1335, 1998. 22

[101] S. I. Rabben, N. Stergiopulos, L. R. Hellevik, O. A. Smiseth, S. Slordahl, S. Urheim, and B. Angelsen. An ultrasound-based method for determining pulse wave velocity in superficial arteries. *J Biomech*, 37(10):1615–22, 2004. 22

[102] S. Graf, D. Craiem, and R.L. Armentano. Non invasive assessment of carotid and femoral arterial pressure using B-mode ultrasound diameter waveforms. In *34th Annual International Conference of the IEEE EMBS*, 2012. 22

[103] F. Scalzo, S. Asgari, S. Kim, M. Bergsneider, and X. Hu. Robust peak recognition in intracranial pressure signals. *Biomed Eng Online*, 9:61, 2010. 22

[104] X. Hu, P. Xu, F. Scalzo, P. Vespa, and M. Bergsneider. Morphological clustering and analysis of continuous intracranial pressure. *IEEE Transactions On Biomedical Engineering*, 56(3):696–705, 2009. 22

[105] F. Scalzo, S. Asgari, S. Kim, M. Bergsneider, and X. Hu. Bayesian tracking of intracranial pressure signal morphology. *Artif Intell Med*, 54(2):115–23, 2012. 22

[106] G. M. London, J. Blacher, B. Pannier, A. P. Guerin, S. J. Marchais, and M. E. Safar. Arterial wave reflections and survival in end-stage renal failure. *Hypertension*, 38(3):434–438, 2001. 23

[107] H. Ueda, T. Hayashi, K. Tsumura, K. Yoshimaru, Y. Nakayama, and J. Yoshikawa. The timing of the reflected wave in the ascending aortic pressure predicts restenosis after coronary stent placement. *Hypertension Research*, 27(8):535–540, 2004. 23

[108] J. A. Chirinos, J. P. Zambrano, S. Chakko, A. Veerani, A. Schob, H. J. Willens, G. Perez, and A. J. Mendez. Aortic pressure augmentation predicts adverse cardiovascular events in patients with established coronary artery disease. *Hypertension*, 45(5):980–5, 2005. 23

[109] T. Weber, J. Auer, M.F. ORourke, E. Kvas, E. Lassnig, G. Lamm, N. Stark, M. Rammer, and E. Eber. Increased arterial wave reections predict severe cardiovascular events in patients undergoing percutaneous coronary interventions. *European Heart Journal*, 26:26572663, 2005. 23

[110] A. M. Dart et al. Brachial blood pressure but not carotid arterial waveforms predict cardiovascular events in elderly female hypertensives. *Hypertension*, 47(4):785–90, 2006. 23

[111] A. Covic, N. Mardare, P. Gusbeth-Tatomir, O. Prisada, R. Sascau, and D. J. Goldsmith. Arterial wave reflections and mortality in haemodialysis patients–only relevant in elderly, cardiovascularly compromised? *Nephrol Dial Transplant*, 21(10):2859–66, 2006. 23

[112] B. Williams, P. S. Lacy, S. M. Thom, K. Cruickshank, A. Stanton, D. Collier, A. D. Hughes, H. Thurston, and M. O'Rourke. Differential impact of blood pressure-lowering drugs on central aortic pressure and clinical outcomes: principal results of the conduit artery function evaluation (CAFE) study. *Circulation*, 113(9):1213–25, 2006. 23

[113] C. Manisty, J. Mayet, R. J. Tapp, K. H. Parker, P. Sever, N. R. Poulter, S. A. Thom, and A. D. Hughes. Wave reflection predicts cardiovascular events in hypertensive individuals independent of blood pressure and other cardiovascular risk factors: an ASCOT (Anglo-Scandinavian Cardiac Outcome Trial) substudy. *J Am Coll Cardiol*, 56(1):24–30, 2010. 23, 154

[114] K. L. Wang, H. M. Cheng, S. H. Sung, S. Y. Chuang, C. H. Li, H. A. Spurgeon, C. T. Ting, S. S. Najjar, E. G. Lakatta, F. C. Yin, P. Chou, and C. H. Chen. Wave reflection and arterial

stiffness in the prediction of 15-year all-cause and cardiovascular mortalities: a community-based study. *Hypertension*, 55(3):799–805, 2010. 23

[115] J. H. Janner, N. S. Godtfredsen, S. Ladelund, J. Vestbo, and E. Prescott. High aortic augmentation index predicts mortality and cardiovascular events in men from a general population, but not in women. *Eur J Prev Cardiol.*, 30(5), 2012. 23

[116] P. S. Sever et al. Rationale, design, methods and baseline demography of participants of the Anglo-Scandinavian Cardiac Outcomes Trial. *J Hypertens.*, 19(6):1139–47, 2001. 23

[117] N. A. Jatoi, A. Mahmud, K. Bennett, and J. Feely. Assessment of arterial stiffness in hypertension: comparison of oscillometric (Arteriograph), piezoelectronic (Complior) and tonometric (Sphygmocor) techniques. *J Hypertens*, 27(11):2186–91, 2009. 24, 26, 27

[118] Complior website. http://www.complior.com/products/, 2012. 24

[119] R. Asmar, A. Benetos, J. Topouchian, P. Laurent, P. Pannier, A. Brisac, R. Target, and B. I. Levy. Assessment of arterial distensibility by automatic pulse wave velocity measurement. *Hypertension*, 26:485–490, 1995. 24, 29

[120] B. Sztrymf, F. Jacobs, D. Chemla, C. Richard, and S. C. Millasseau. Validation of the new complior sensor to record pressure signals non-invasively. *J Clin Monit Comput*, 2013. 24, 25

[121] Sphygmocor website. http://www.atcormedical.com/sphygmocor.html, 2012. 25

[122] R. Kelly, C. Hayward, J. Ganis, J. Daley, A. Avolio, and M. O'Rourke. Noninvasive registration of the arterial pressure pulse waveform using high-fidelity applanation tonometry. *J Vasc Med Biol.*, 3:142–9, 1989. 25

[123] S.A. Hope, I.T. Meredith, and Cameron J.D. Effect of non-invasive calibration of radial waveforms on error in transfer-function-derived central aortic waveform characteristics. *Clin Sci (Lond).*, 107(2):205–11, 2004. 25

[124] H. Smulyan, D. S. Siddiqui, R. J. Carlson, G. M. London, and M. E. Safar. Clinical utility of aortic pulses and pressures calculated from applanated radial-artery pulses. *Hypertension*, 42(2):150–5, 2003. 25

[125] G. C. Cloud, C. Rajkumar, J. Kooner, J. Cooke, and C. J. Bulpitt. Estimation of central aortic pressure by Sphygmocor requires intra-arterial peripheral pressures. *Clinical Science*, 105:219–225, 2003. 25

[126] S. Dyer, M. Bevan, and J. Hutchinson. Peripheral arterial tonometry with ascending aortic waveform analysis using the Sphygmocor system. Technical report, Commonwealth of Australia, 2006. 25

[127] Y. Zhang, D. Agnoletti, M. E. Safar, J. G. Wang, J. Topouchian, Y. Xu, A. D. Protogerou, and J. Blacher. Comparison study of central blood pressure and wave reflection obtained from tonometry-based devices. *Am J Hypertens*, 26(1):34–41, 2013. 26

[128] L. Joly, C. Perret-Guillaume, A. Kearney-Schwartz, P. Salvi, D. Mandry, P. Y. Marie, G. Karcher, P. Rossignol, F. Zannad, and A. Benetos. Pulse wave velocity assessment by external noninvasive devices and phase-contrast magnetic resonance imaging in the obese. *Hypertension*, 54(2):421–6, 2009. 26

[129] P. Salvi, G. Lio, C. Labat, E. Ricci, B. Pannier, and A. Benetos. Validation of a new non-invasive portable tonometer for determining arterial pressure wave and pulse wave velocity: the PulsePen device. *J Hypertens.*, 22(12):2285–93, 2004. 26, 29

[130] P. Salvi, E. Magnani, F. Valbusa, D. Agnoletti, C. Alecu, L. Joly, and A. Benetos. Comparative study of methodologies for pulse wave velocity estimation. *J Hum Hypertens*, 22(10):669–77, 2008. 27

[131] S. L. Magda, A. O. Ciobanu, M. Florescu, and D. Vinereanu. Comparative reproducibility of the noninvasive ultrasound methods for the assessment of vascular function. *Heart Vessels*, 28(2):143–50, 2013. 27, 29

[132] A. Margulescu, M. Cinteza, and D. Vinereanu. Reproducibility in echocardiography: clinical significance, assessment, and comparison with other imaging methods. *Medica: A Journal of Clinical Medicine*, 1(2), 2006. 27

[133] M. Frimodt-Moller, A. H. Nielsen, A. L. Kamper, and S. Strandgaard. Reproducibility of pulse-wave analysis and pulse-wave velocity determination in chronic kidney disease. *Nephrol Dial Transplant*, 23(2):594–600, 2008. 29, 34, 147

[134] A. Siebenhofer, C. Kemp, A. Sutton, and B. Williams. The reproducibility of central aortic blood pressure measurements in healthy subjects using applanation tonometry and sphygmocardiography. *J Hum Hypertens.*, 13(9):625–9, 1999. 29, 147

[135] M. Crilly, C. Coch, M. Bruce, H. Clark, and D. Williams. Indices of cardiovascular function derived from peripheral pulse wave analysis using radial applanation tonometry: a measurement repeatability study. *Vasc Med*, 12(3):189–97, 2007. 29, 147

[136] S. Wassertheurer, J. Kropf, T. Weber, M. van der Giet, J. Baulmann, M. Ammer, B. Hametner, C. C. Mayer, B. Eber, and D. Magometschnigg. A new oscillometric method for pulse wave analysis: comparison with a common tonometric method. *J Hum Hypertens*, 24(8):498–504, 2010. 29

[137] S. Mora, L. R. Yanek, T. F. Moy, M. D. Fallin, L. C. Becker, and D. M. Becker. Interaction of body mass index and Framingham risk score in predicting incident coronary disease in families. *Circulation*, 111(15):1871–6, 2005. 30

[138] J. Janner, N. S Godtfredsen, S. Ladelund, J. Vestbo, and E. Prescott. Aortic augmentation index: reference values in a large unselected population by means of the Sphygmocor device. *American journal of hypertension*, 23(2):180–5, 2010. 31

[139] W. Wojciechowska, J.A. Staessen, T. Nawrot, M. Cwynar, J. Seidlerová, K. Stolarz, J. Gsowski, M. Tichá, T. Richart, L. Thijs, T. Grodzicki, K. Kawecka-Jaszcz, and J. Filipovsk. Reference values in white Europeans for the arterial pulse wave recorded by means of the Sphygmocor device. *Hypertension Research*, 29:475–483, 2006. 32

[140] G. F. Mitchell, H. Parise, E. J. Benjamin, M. G. Larson, M. J. Keyes, J. A. Vita, R. S. Vasan, and D. Levy. Changes in arterial stiffness and wave reflection with advancing age in healthy men and women: the Framingham Heart Study. *Hypertension*, 43(6):1239–45, 2004. 32, 34, 116

[141] Y. Li, J. A. Staessen, L. H. Li, Q. F. Huang, L. Lu, and J. G. Wang. Reference values for the arterial pulse wave in Chinese. *Am J Hypertens*, 21(6):668–73, 2008. 31, 32

[142] J.W. Chung, Y.S. Lee, J.H. Kim, M.J. Seong, S.Y. Kim, J.B. Lee, J.K. Ryu, J.Y. Choi, K.S. Kim, S.G. Chang, G.H. Lee, and S.H. Kim. Reference values for the augmentation index and pulse pressure in apparently healthy Korean subjects. *Korean circulation journal*, 40(4):165–71, 2010. 31, 32

[143] C. P. Shiburi, J. A. Staessen, M. Maseko, W. Wojciechowska, L. Thijs, L. M. Van Bortel, A. J. Woodiwiss, and G. R. Norton. Reference values for Sphygmocor measurements in South Africans of African ancestry. *Am J Hypertens*, 19(1):40–6, 2006. 31, 32

[144] C. M. McEniery, Yasmin, I. R. Hall, A. Qasem, I. B. Wilkinson, and J. R. Cockcroft. Normal vascular aging: differential effects on wave reflection and aortic pulse wave velocity: the Anglo-Cardiff Collaborative Trial (ACCT). *J Am Coll Cardiol*, 46(9):1753–60, 2005. 33, 125

[145] M. F. Elias, G. A. Dore, A. Davey, W. P. Abhayaratna, A. L. Goodell, and M. L. Robbins. Norms and reference values for pulse wave velocity: one size does not fit all. *The Journal of Bioscience and Medicine*, 1(1):1–10, 2011. 33

[146] R. E. W. Kavey, S. R. Daniels, R. M. Lauer, D. L. A. Atkins, L. L. Hayman, and K. Taubert. American heart association guidelines for primary prevention of atherosclerotic cardiovascular disease beginning in childhood. *Circulation*, 107(11):1562–1566, 2003. 34

[147] A. Beswick and P. Brindle. Risk scoring in the assessment of cardiovascular risk. *Curr Opin Lipidol*, 17:375–386, 2006. 34

[148] M. D. Whitfield, M. Gillett, M. Holmes, and E. Ogden. Predicting the impact of population level risk reduction in cardio-vascular disease and stroke on acute hospital admission rates over a 5 year period–a pilot study. *Public Health*, 120(12):1140–8, 2006. 34

[149] M. T. Cooney, A. L. Dudina, and I. M. Graham. Value and limitations of existing scores for the assessment of cardiovascular risk: a review for clinicians. *J Am Coll Cardiol*, 54(14):1209–27, 2009. 35, 100

[150] M. G. Tsipouras, T. P. Exarchos, D. I. Fotiadis, A. Kotsia, K. Vakalis, K. K. Naka, and L. K. Michalis. Automated diagnosis of coronary artery disease based on data mining and fuzzy modeling. *IEEE Transactions On Information Technology In Biomedicine*, 12(4):447–458, 2008. 35

[151] A. Jovic and N. Bogunovic. Electrocardiogram analysis using a combination of statistical, geometric, and nonlinear heart rate variability features. *Artificial Intelligence in Medicine*, 51:175–186, 2011. 35

[152] Q. Y. Lee, G. S. H. Chan, S. J. Redmond, P. M. Middleton, E. Steel, P. Malouf, C. Critoph andG. Flynn, E. O'Lone, and N. H. Lovell. Multivariate classification of systemic vascular resistance using photoplethysmography. *Physiol. Meas.*, 32:1117–1132, 2011. 35

[153] B. M. Asl, S. K. Setarehdan, and M. Mohebbi. Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal. *Artif Intell Med*, 44(1):51–64, 2008. 35

[154] S. L. Ting, C. C. Shym, S. K. Kwork, A. H. C. Tsang, and W. B. Lee. Data mining in biomedicine: current applications and further directions for research. *Journal of Software Engineering and Applications*, 2(3):150–159, 2009. 35

[155] K. B. Wagholikar, V. Sundararajan, and A. W. Deshpande. Modeling paradigms for medical diagnostic decision support: a survey and future directions. *J Med Syst*, 36(5):3029–49, 2012. 35

[156] L. I. Kuncheva, J. C. Bezdek, and R. P.W. Duin. Decision templates for multiple classifier fusion: an experimental comparison. *Pattern Recognition Letters*, 34(2), 2001. 36

[157] D. Ruta and B. Gabrys. An overview of classifier fusion methods. *Computing and Information Systems*, 7, 2000. 36

[158] J. Han and M. Kamber. *Data mining. Concepts and techniques*. Elsevier, San Francisco, CA, 2006. 36, 94, 96, 99, 101

[159] V. Melnykov and G. Shen. Clustering through empirical likelihood ratio. *Computational Statistics & Data Analysis*, 2013. 36

[160] A. S. Shah, L. M. Dolan, Z. Gao, T. R. Kimball, and E. M. Urbina. Clustering of risk factors: a simple method of detecting cardiovascular disease in youth. *Pediatrics*, 127(2):e312–8, 2011. 36

[161] H. H. Haseena, A. T. Mathew, and J. K. Paul. Fuzzy clustered probabilistic and multi layered feed forward neural networks for electrocardiogram arrhythmia classification. *J Med Syst*, 35(2):179–88, 2011. 36

[162] J. Namayanja and V. P. Janeja. An assessment of patient behavior over timeperiods: a case study of managing Type 2 diabetes through blood glucose readings and insulin doses. *Journal of Medical Systems*, 36(S1):65–80, 2012. 36

[163] H. C. Pereira, J. Maldonado, T. Pereira, M. Contente, J. Cardoso, V. Almeida, and C. Correia. A novel and low cost acoustic based probe for local pulse wave velocity estimation. In *6th International Conference on Biomedical Electronics and Devices*, pages 79–88, 2013. 38

[164] P. Santos. Photoplethysmographic logger with contact force and hydrostatic pressure monitoring. Master's thesis, University of Coimbra, 2012. 38, 42

[165] T. Pereira, P. Vaz, T. Oliveira, I. Santos, H. C. Pereira, V. Almeida, C. Correia, and J. Cardoso. Empirical mode decomposition for self-mixing doppler signals of hemodynamic optical probes. *Physiol Meas*, 34(3):377–90, 2013. 38

[166] V. G. Almeida, H. C. Pereira, T. Pereira, E. Figueiras, E. Borges, J. M. R. Cardoso, and C. Correia. Piezoelectric probe for pressure waveform estimation in flexible tubes and its applications to the cardiovascular system. *Sensors and Actuators A*, 169:217–226, 2011. 38, 75

[167] C. Leitão, L. Bilro, N. Alberto, P. Antunes, H. Lima, P. S. André, R. Nogueira, and J. L. Pinto. Feasibility studies of bragg probe for noninvasive carotid pulse waveform assessment. *Journal of Biomedical Optics*, 18(1):017006–1–017006–2, 2013. 38

[168] P. Vaz. Synchronization algorithm. Forthcoming, 2013. 43

[169] P. Santos. Hemodynamic database. User guide, 2013. 44

[170] J. Karki. Signal conditioning piezoelectric sensors - SLOA033A. Technical report, Texas Instruments, 2000. 49

[171] Piezoelectric sound components - P37E17.pdf 02.3.6. Datasheet, Murata, 2001. 48, 49

[172] J. McLaughlin, M. McNeill, B. Braun, and P. D. McCormack. Piezoelectric sensor determination of arterial pulse wave velocity. *Physiol Meas*, 24:693–702, 2003. 49

[173] L. Robinson. Reduction of baseline shift in pulse-amplitude measurements. *Rev. Sci. Instr.*, 32:1057, 1961. 52

[174] Wide bandwidth operational transconductance amplifier OTA and buffer - SBOS331C. Datasheet, Texas Instruments, 2008. 53

[175] K.H. Chon, S. Dash, and Kihwan Ju. Estimation of respiratory rate from photoplethysmogram data using time. *Biomedical Engineering, IEEE Transactions on*, 56(8):2054–2063, 2009. 54

[176] V. Klig. Biomedical applications of microprocessors. *Proceedings of the IEEE*, 66(2):151–161, 1978. 56

[177] B. N Li, M. C. Dong, M. I. Vai, and M. P. Un. An embedded medical advisory system for mobile cardiovascular monitoring devices. In *Biomedical Circuits and Systems, 2004 IEEE International Workshop on*, pages 1–1–4, 2004. 56

[178] J. Germano, R. Ramalho, and L. Sousa. On the design of distributed autonomous embedded systems for biomedical applications. In *Pervasive Health*, pages 1–8, 2009. 56

[179] B. Buteler. The relation of systolic upstroke time and pulse pressure in aortic stenosis. *Br Heart J.*, 24(5):657–660, 1962. 57

[180] L. Kerhuel. Embedded target for dsPIC. http://www.kerhuel.eu/, 2010. 59, 161

[181] V. Almeida, T. Pereira, E. Borges, E. Figueiras, C. Correia J. Cardoso, H.C. Pereira, J. L. Malaquias, and J. B. Simes. Synthesized cardiac waveform in the evaluation of augmentation index algorithms. In *3rd International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2010)*, 2010. 65, 68, 82, 85

[182] J.P. Murgo, N. Westerhof, J. P. Giolma, and S.A. Altobelli. Aortic input impedance in normal man: relationship to pressure wave forms, circulation. *Circulation*, 62:105–116, 1980. 65, 91

[183] V. G. Almeida, P. Santos, E. Figueiras, E. Borges, T. Pereira, J. Cardoso, C. Correia, and H. C. Pereira. Hemodynamic features extraction from a new arterial pressure waveform probe. In *Proceeding of the 4th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2011)*, Rome, Italy, January 2011. 75

[184] J. Borba. Validation of non-invasive electromechanical sensors for cardiac monitoring: clinical trials and implementation of data mining techniques. Master's thesis, University of Coimbra, 2012. 77, 139

[185] AXIOM Sensis XP. http://usa.healthcare.siemens.com/interventional-cardiology/card-angiography/axiom-sensis-xp, 2013. 77

[186] A. Donelli, J. R. C. Jansen, B. Hoeksel, P. Pedeferri, R. Hanania, J. Bovelander, F. Maisano, A. Castiglioni, O. Alfieri, and J. J. Schreuder. Performande of a real-time dicrotic notch detection and prediction algorithm in arrhythmic human aortic pressure signals. *J Clin Monit*, 17:181–185, 2002. 85

[187] M. Oppenheim and D. F. Sittig. An innovative dicrotic notch detection algorithm which combines rule-based logic with digital signal processing techniques. *Computers and biomedical research*, 28:154–170, 1995. 85

[188] M. Melis, U. Morbiducci, E. R. Rietzschel, M. Buyzere, A. Qasem, L. Bortel, T. Claessens, F. M. Montevecchi, A. Avolio, and P. Segers. Blood pressure waveform analysis by means of wavelet transform. *Medical & Biological Engineering Computing*, 47(2):165–173, 2008. 85

[189] L. A. Bortolotto, J. Blacher, T. Kondo, K. Takazawa, and M. E. Safar. Assessment of vascular aging and atherosclerosis in hypertensive subjects: second derivative of photoplethysmogram versus pulse wave velocity. *AJH*, 13:65–171, 2000. 85

[190] V. Almeida. Hemodynamic parameters assessment. Master's thesis, University of Coimbra, 2009. 85

[191] P. H. Tsui, L. Y. Lin, C. C. Chang, J. J. Hwang, J. J. Lin, C. C. Chu, C. N. Chen, and K. J. Chang. Arterial pulse waveform analysis by the probability distribution of amplitude. *Physiol Meas*, 28(8):803–12, 2007. 85

[192] B. N. Li, M. C. Dong, and M. I. Vai. On an automatic delineator for arterial blood pressure waveforms. *Biomedical Signal Processing and Control*, 5(1):76–81, 2010. 87

[193] W. Zong, T. Heldt, G. B. Moody, and R. G. Mark. An open-source algorithm to detect onset of arterial blood. *Computers in Cardiology*, 30:259–262, 2003. 87

[194] I. H. Witten, E. Frank, and M. A. Hall. *Data Mining: practical Machine Learning Tools and Techniques*. 2nd edn. Morgan Kaufmann, San Francisco, 2005. 95, 96, 101

[195] J. R. Quinlan. *C4.5: programs for Machine Learning*. Morgan Kaufmann Publishers, San Mateo, California, 1993. 95

[196] L. Breiman. Random forests. *Machine Learning*, 45:5–32, 2001. 96

[197] N. Gayatri, Nickolas S, A. V. Reddy, and R. Chitra. Performance analysis of datamining algorithms for software quality prediction. In *International Conference on Advances in Recent Technologies in Communication and Computing*, pages 393–395. Curran Associates, Inc., 2009. 96

[198] H. A. Nguyen and D. Choi. Application of data mining to network intrusion detection: classifier selection model. *Springer-Verlag Berlin Heidelberg*, 2008. 96

[199] S. Haykin. *Neural Networks - A Comprehensive Foundation*. Pearson education, India, 1998. 97

[200] L. Rokach. Ensemble-based classifiers. *Artificial Intelligence Review*, 33(1-2):1–39, 2009. 97

[201] Y. H. Hu, F. Wu, C. L. Lo, and C. T. Tai. Predicting warfarin dosage from clinical data: a supervised learning approach. *Artif Intell Med*, 2012. 98

[202] I. U. Haq, L. E. Ramsay, W.W. Yeo, P. R. Jackson, and E. J. Wallis. Is the Framingham risk function valid for northern European populations? a comparison of methods for estimating absolute coronary risk in high risk men. *Heart Vessels*, 81:40–46, 1999. 100, 114, 126

[203] J. Foussier, P. Fonseca, X. Long, and S. Leonhardt. Automatic feature selection for sleep/wake classication with small data sets. In Mireya Fernández Chimeno, Jordi Solé-Casals, Ana Fred, , and Hugo Gamboa, editors, *6th International Conference on Biomedical Electronics and Devices*, volume Bioinformatics 2013. SCITEPRESS Science and Technology Publications, 2013. 102

[204] B. G. Song, J. B. Park, S. J. Cho, S. Y. Lee, J. H. Kim, S. M. Choi, J. H. Park, Y. H. Park, J. O. Choi, S. C. Lee, and S. W. Park. Pulse wave velocity is more closely associated with cardiovascular risk than augmentation index in the relatively low-risk population. *Heart Vessels*, 24(6):413–8, 2009. 116

[205] M. Woodward, P. Brindle, and H. Tunstall-Pedoe. Adding social deprivation and family history to cardiovascular risk assessment: the ASSIGN score from the Scottish Heart Health Extended Cohort (SHHEC). *Heart*, 93(2):172–6, 2007. 126

[206] J. Pallant. *SPSS Survival Manual*. Allen Unwin, Australia, second edition, 2005. 131, 132, 133, 134

[207] H. M. Kaltenbach. *A Concise Guide to Statistics*. Springer, German, 2012. 131

[208] J. Cohen. *Statistical power analysis for the behavioral sciences*. Erlbaum, New York, second edition, 1988. 132

[209] J. Hauke and T. Kossowski. Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data. *Quaestiones Geographicae*, 30(2), 2011. 133

[210] O. Vriz, E. Bossone, M. Bettio, D. Pavan, S. Carerj, and F. Antonini-Canterin. Carotid artery stiffness and diastolic function in subjects without known cardiovascular disease. *J Am Soc Echocardiogr*, 24(8):915–21, 2011. 154

[211] S. B. Kotsiantis. Supervised machine learning: a review of classification techniques. *Informatica*, 31:249–268, 2007. 155